Simulation of Colour Evolution in QCD Scattering Processes

A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy in the Faculty of Engineering and Physical Sciences

2013

Alexander Anthony Schofield

Particle Physics Group

School of Physics and Astronomy

Contents

1 Quantum Chromodynamics **11**

1.1 Introduction . . . 11
1.2 QCD Lagrangian . . . 12
1.3 The strong coupling . . . 13
1.4 Perturbation Theory . . . 14
1.4.1 Soft and Collinear Limits . . . 16
1.4.2 Resummation . . . 17
1.4.3 Eikonal limit . . . 18
1.5 Colour evolution . . . 19
1.5.1 Leading colour . . . 21
1.6 Colour states and evolution . . . 22
1.7 Colour emission and exchanges . . . 24
1.8 Colour bases . . . 25
1.8.1 Orthonormal basis . . . 25
1.8.2 Colour flow basis . . . 27
1.8.3 Relating the bases . . . 29
1.8.4 Special Case: $gg \to gg$ . . . 29
1.9 Numerical Simulation of QCD . . . 31
1.9.1 Hard Process . . . 32
1.9.2 Parton Shower . . . 33
1.9.3 Multiple Interactions . . . 37
1.9.4 Hadronization . . . 38
1.10 Conclusion . . . 38

2 Jet Vetoes and Herwig++ **42**

2.1 Introduction . . . 42
2.2 Gaps Between Jets - Analytical Approach . . . 42
2.3 Parton shower differences . . . 46
2.4 Analytical Modification . . . 49
2.4.1 Analytical "parton shower" model . . . 49
2.4.2 Partonic Fractions . . . 52
2.4.3 Energy-momentum conservation . . . 56
2.5 Parton shower modification . . . 57
2.6 Results . . . 59
2.7 Conclusion . . . 65

3 Tuning for Wide-Angle Radiation 69
3.1 Introduction . . . 69
3.2 General Tuning Considerations . . . 70
3.2.1 Models within the event generator . . . 70
3.2.2 Selection of tuning parameters . . . 71
3.2.3 Statistics . . . 72
3.2.4 Tuning procedure . . . 74
3.3 Tuning for wide-angle radiation . . . 74
3.3.1 Summary of changes . . . 74
3.3.2 Selection of relevant tuning parameters . . . 75
3.3.3 Interpretation of parameters . . . 77
3.3.4 Parameter ranges . . . 79
3.3.5 Selection of relevant analyses . . . 81
3.3.6 Statistics . . . 85
3.3.7 Tuning approach . . . 87
3.3.8 Improving with additional data . . . 87
3.4 Tunes . . . 87
3.4.1 Tune 1: Gaps Between Jets with no bins excluded 88
3.4.2 Tune 2: Gaps Between Jets with higher bins excluded 88
3.4.3 Tune 3: Gaps Between Jets without forward and backward observables . . . 89
3.4.4 Tune 4: Gaps Between Jets and Dijets . . . 89
3.4.5 Tune 5: All LHC analyses with some bins excluded 89
3.4.6 Tune 6: Pure e^+e^- Tune . . . 89
3.4.7 Tune 7: e^+e^- and LHC Tune . . . 90
3.5 Results . . . 90
3.5.1 Correlations . . . 90
3.5.2 Tune values . . . 91
3.5.3 Resulting plots . . . 93
3.6 Conclusion . . . 97

4 Treatment of Sub-Leading Colour in Event Generators 104
4.1 Introduction . . . 104
4.2 Sources of Sub-Leading Colour . . . 105
4.2.1 Colour in the eikonal approach . . . 105
4.2.2 Other sources of Sub-Leading Colour . . . 108
4.2.3 Relative sizes of effects . . . 109
4.3 Other sub-leading colour implementations . . . 109
4.3.1 Plätzer-Sjödahl Approach . . . 109

4.3.2 Nagy-Soper Approach . . . 111
4.4 Colour Exchanges: General Consideration . . . 113
4.4.1 Competing processes . . . 113
4.4.2 Concept of Evolution Time . . . 114
4.4.3 Variable colour structure . . . 117
4.5 Colour Exchanges: Phenomenological Model . . . 118
4.5.1 Overview . . . 118
4.5.2 Input stage . . . 121
4.5.3 Initial Colour Selection . . . 121
4.5.4 The Main Loop . . . 122
4.5.5 Choice of Amplitude . . . 123
4.5.6 Choice of Interaction . . . 124
4.5.7 Splitting . . . 126
4.5.8 Colour Exchange . . . 129
4.5.9 End of the loop . . . 130
4.5.10 End of the shower . . . 131
4.5.11 Example event . . . 131
4.6 Implementation in regular parton showers . . . 133
4.7 Conclusion . . . 135

5 Final Conclusion **138**

Word count: 35277

List of Figures

1 Gluon interaction vertices. 13
2 The colour line decomposition of an exchanged gluon. . . 19
3 Anti-symmetric three gluon coupling. 20
4 Kinematical cuts for one gluon exchange. 25
5 The three orthogonal basis states for $qg \to qg$. 26
6 Symmetric three gluon coupling. 27
7 The three colour flow basis states for $qg \to qg$. 29
8 A schematic representation of an event generator. 32
9 The two possible colour line assignments in $g \to gg$. 37
10 Dijet production with a rapidity gap. 44
11 The colour diagram for the $t_i \cdot t_j$ terms. 46
12 Colour line connections for an exchanged gluon. 48
13 Default radiation patterns for Herwig++. 49
14 Colour line representation of $g \to gg$ splitting. 49
15 Partonic fractions for Y=1, Q=500 GeV. 54
16 Partonic fractions for Y=5, Q=500 GeV. 55
17 Effects of partial energy-momentum conservation. 58
18 New radiation patterns for Herwig++. 59
19 The gap fraction as a function of Q_0. 61
20 Colour evolution modified to unmodified ratio. 62
21 Parton shower to analytical ratio. 62
22 Hadronization ratio. 63
23 Overall gap cross section. 64
24 Gap fractions for `AlphaMZ`. 80
25 Gap fractions for `cutoffKinScale`. 81
26 Gap fractions for `PSplitLight`. 82
27 Gap fractions for `ClPowLight`. 83
28 Gap fractions for `ClMaxLight`. 84
29 Gap fraction statistics comparison. 86
30 Tune 4 correlations. 91
31 Tune 6 correlations. 92
32 Tune 7 correlations. 93
33 Gap fraction for hardest jets. 96
34 Gap fraction for forward and backwards jets. 97
35 Dijet azimuthal correlations. 98
36 Charged hadron multiplicity as a function of rapidity. . . 99
37 Final state parton shower evolution example. 115

38 Final state dipole shower evolution example. 116
39 The overall flow of the algorithm. 119
40 An example of a junction for $gg \to gg$. 121
41 Main event loop. 122
42 A colour exchange resulting in the change of a junction. . 130
43 A simplified example of an event within the algorithm. . 132

List of Tables

1 Default parameter values and variance. 79
2 Observable weights. 88
3 Parameter values for tunes 1-4. 94
4 Parameter values for tunes 5-7. 95
5 An example of a junction values for $gg \to gg$. 120
6 Kinematics for splittings. 128
7 Colour line assignments. 129

Abstract

The University of Manchester
Alexander Anthony Schofield
A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy in the Faculty of Engineering and Physical Sciences
Simulation of Colour Evolution in QCD Scattering Processes
26/09/2013

We investigate the effects of colour evolution in QCD scattering processes and how these can be implemented in both analytical and numerical approaches. We split this in to four parts where each part is given in one chapter.

In the first chapter we give a brief summary of the important aspects of QCD which are needed as a basis for the rest of the investigation. In addition to this, we describe different sets of formalisms for handling colour within interactions. We then give a brief review of the components of a Monte-Carlo event generator.

In the second chapter we review previous work by the author on jet vetoes and their implementation in the Monte-Carlo event generator Herwig++. We describe the analytical method for studying jet vetoes and then discuss the differences between this method and that which is used in the original parton shower of Herwig++. Once this is done we make changes to both the analytical approach and Herwig++ in order to investigate these differences. We then show the results for an improved parton shower as a result of this investigation.

In the third chapter we consider the effects of tuning the parameters within Herwig++. We investigate what parameters are likely to have the most changes to observables given the modifications made in the previous chapter. We then produce seven tunes to different sets of observables and discuss said tunes.

In the fourth and final chapter we discuss the effects of sub-leading colour within the analytical approach and in a potential numerical setup. We discuss a set of potential algorithms for implementing sub-leading colour within a standalone parton shower.

Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright Statement

i The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the Copyright) and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

iii The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the Intellectual Property) and any reproductions of copyright works in the thesis, for example graphs and tables (Reproductions), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may

take place is available in the University IP Policy (see http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=487), in any relevant Thesis restriction declarations deposited in the University Library, The University Librarys regulations (see http://www.manchester.ac.uk/library/aboutus/regulations) and in The Universitys policy on Presentation of Theses.

Acknowledgements

First of all I would like to thank my supervisor Mike Seymour for all his support throughout my postgraduate career. His patient guidance and encouragement were always there when I needed it. Our discussions greatly helped with my understanding of my research as well as physics in general.

Next I would like to thank the University of Manchester Particle Physics Group for all of their help. The pleasant and helpful atmosphere throughout the group was a great motivating factor and I had many enjoyable times. In particular, I would like to thank Sabah Salih for his extensive help with the computer systems.

I would also like to thank the members of the Herwig++ collaboration for their help with implementing physics within event generators and methods for producing events efficiently. I would especially like to thank Andrzej Siodmok for his help with the tuning procedure.

Finally I would like to thank my parents Ann and Tony Schofield for their constant support throughout my studies.

1 Quantum Chromodynamics

1.1 Introduction

In order to interpret the results at experiments such as the Large Hadron Collider and Tevatron we must have a thorough understanding of the underlying theory which contains the relevant interactions that are occuring. Establishing a connection between the theoretical predictions and experimental observables is the goal of particle physics phenomenology.

To make this connection there are two different approaches. The first is the analytical calculation, where the theory is applied to calculating a specific observable. While it is in theory possible to calculate any observable to an arbitrary level of precision, once the final state of said observable becomes complex the analytical calculation can become unwieldy. One area where the analytical calculations are especially effective is the study of high energy jets. When calculating an observable with a large number of final-state particles approximations are normally used to simplify the computation.

The second is the numerical approach, where a set of algorithms built on physical principles or models is used to study a wide array of phenomena. There are various numerical approaches, such as lattice QCD or Monte-Carlo event generators. In this chapter and those that follow we will focus on the latter. These event generators excel at simulating events with large amounts of soft radiation, which is what we generally observe from QCD interactions. Since these numerical approaches need to have a well defined algorithm it is not possible, with our current understanding of QCD, to completely contain the analytical approach within our event generators.

1.2 QCD Lagrangian

Quantum chromodynamics is the theory of the strong force that dictates the interactions between quarks and gluons. These quarks and gluons form the hadrons which are observed in nature. The symmetries of the force can be expressed in terms of a Lagrangian as

$$L_{QCD} = \bar{\psi}_i(i\gamma^\mu(D_\mu)_{ij} - m\delta_{ij})\psi_j - \tfrac{1}{4}G^A_{\mu\nu}G^{\mu\nu}_A. \tag{1.1}$$

Here, and throughout this chapter, we use the notation that any repeated index is summed over. The lower case indices i and j run from 1 to N_c, where N_c is the number of colours, and the upper case indices A run from 1 to $N_c^2 - 1$. The first term in the Lagrangian contains the dynamics of the quark fields. It differs from QED only in the fact that there are now colour indices and the covariant derivative contains a colour mixing matrix. The covariant derivative $(D_\mu)_{ij}$ for QCD is given by

$$(D_\mu)_{ij} = \delta_{ij}\partial_\mu - ig_s A^A_\mu T^A_{ij}. \tag{1.2}$$

The second term in the Lagrangian contains the dynamics of the gluon fields. These again differ in the fact that they carry a colour index for the gluons. The field strength tensor $G^A_{\mu\nu}$ for QCD is

$$G^A_{\mu\nu} = \partial_\mu A^A_\nu - \partial_\nu A^A_\mu - g_s f^{ABC} A^B_\mu A^C_\nu. \tag{1.3}$$

From this expression we can see that compared to QED there is an important difference here. When we combine the two field strength tensors we find that gluons are able to have self interactions. That is giving quarks a non-Abelian symmetry over colours requires that the gluons themselves carry colour. This leads to different phenomenological behaviour than QED. The two new interaction vertices in QCD are shown

in Figure 1.

In these equations ψ is the quark field, A^A_μ is the gluon field, g_s is the strong charge, T^A_{ij} and f^{ABC} are generators for the colour group which will be discussed in a later section.

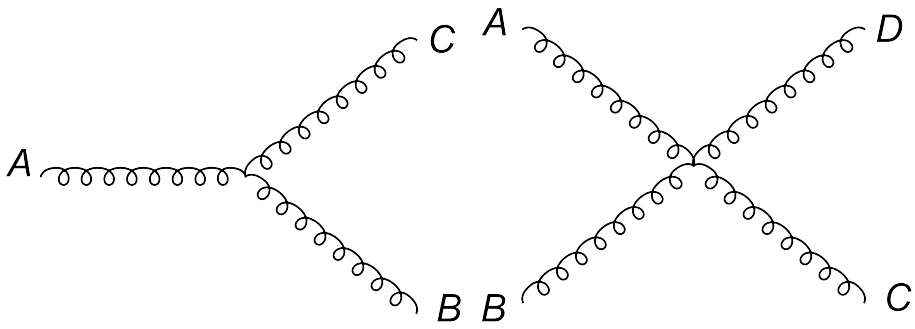

Figure 1: Left: Three gluon interaction vertex. Right: Four gluon interaction vertex.

1.3 The strong coupling

We define a coupling constant for QCD in terms of the colour charge g_s as

$$\alpha_s = \frac{g_s^2}{4\pi}. \tag{1.4}$$

In addition to the colour, QCD also differs in the strength of the coupling. At scales around the mass of the Z boson the strong coupling constant is of order 10^{-1}, which is much greater than the electromagnetic coupling. A 1-loop calculation of the gluon self-coupling allows us to write the coupling at any scale q within the perturbative regime as

$$\alpha_s(q) = \frac{\alpha_s(m_z)}{1 + \alpha_s(m_z) b \log(\frac{q^2}{m_z^2})}, \tag{1.5}$$

where b is given by

$$b = \frac{33 - 2n_f}{12\pi}, \tag{1.6}$$

where n_f is the number of flavours accessible at the given scale. This formula differs from that of QED by the sign in the denominator. As the scale increases the coupling weakens which allows quarks and gluons to behave as free particles. This is the phenomenon known as asymptotic freedom. Conversely at low energies the coupling is stronger, which results in the binding of quarks within hadrons.

Due to the coupling of QCD being generally larger than that of QED we observe much more radiation in the form of gluons than photons when partons are scattered in an interaction. The enhancement to the coupling in the low energy regime results in a large amount of soft radiation being produced.

1.4 Perturbation Theory

Interactions in high energy physics can be thought of in terms of Feynman diagrams. The Feynman diagram with the least number of interactions contributing to a given final state is referred to as the Born, hard or leading order (LO) process. At energies which are low, but still high enough that the α_s allows predictive power, this hard process is approximately what will be observed in nature. However, as the energies are increased, the possibility for additional emissions and colour exchanges occurs. We refer to these additional processes as higher order terms as their cross section will always be proportional to the hard cross section multiplied by at least one power of the coupling constant. These higher order terms can also be expressed in terms of Feynman diagrams. The total cross section $\sigma^T_{ij;kl}$ for a process with initial-state particles i, j and final-state particles k, l can be written as

$$\sigma^T_{ij;kl} = \sigma^H_{ij;kl} \Sigma^{\infty}_{n=0} \alpha_s^n A^{(n)}_{ij;kl} = \sigma^H_{ij;kl} + O(\alpha_s). \tag{1.7}$$

Naively we see that we can expand in powers of the coupling constant in order to achieve an arbitrary amount of precision when calculating the cross section. However the coefficients $A_{ij;kl}^{(n)}$ for higher orders will generally contain many complex diagrams which are difficult to compute. The number of diagrams contributing to a given order grows rapidly as the order increases. These higher order diagrams do not contain only real emission but also virtual colour exchanges.

For inclusive measurements, such as the total cross section, the higher order terms will be proportional only to powers of α_s and constant numerical factors. For exclusive measurements, where certain configurations of final states are rejected by the analysis, the higher order terms will also in general be proportional to large logarithms of the ratio of the hard and soft scales. We define the hard scale as being proportional to some maximum value related to the hard process. One potential choice could be the average transverse momentum of the two hardest jets. The soft scale can be regarded as the minimum value of some parameter at which the analysis still makes sense. While this is highly dependent on the analysis, one possible choice is the maximum transverse momentum allowed for an extra jet between the two hardest jets.

At low values of the soft scale each term can be as important as the last, limiting the predictive power of a truncated perturbative expansion. In order to handle these smaller values for the soft scale another method must be used to regain predictive power. This method, which is known as a resummation, will be explained in a later section.

While Feynman diagrams are useful for visualizing what is occuring in a process they are much more than that. Each line and vertex represents the mathematical factors used in calculating the amplitude for the process displayed.

1.4.1 Soft and Collinear Limits

In QCD most emissions[1] can be thought of as arising from colour connected dipoles. Each of these dipoles will contribute a factor to the overall cross section. For a colour connected pair $\{i, j\}$ with momenta p_i and p_j emitting a gluon with momentum k the contribution is

$$D_{ij} = \Omega_{ij} \frac{p_i \cdot p_j}{(p_i \cdot k)(p_j \cdot k)}, \tag{1.8}$$

where the factor Ω_{ij} is present to indicate that different dipoles can have different colour prefactors. There are clearly singularities in this dipole when k is collinear with either of the two emitting partons or $k \to 0$. After integrating over the additional phase space generated by this emission we still have the problematic singularities.

In order to remove these singularities we need to also consider interactions where an additional gluon is exchanged within the process, rather than emitted. These are referred to as virtual interactions or colour exchanges. These virtual interactions contain loops over internal parton momenta which are divergent at exactly the same point, but the poles they generate are equal and opposite to those of the real emission. When the virtual exchanges are combined with the real emissions the singularities are removed. For inclusive observables what remains will be terms in α_s. For exclusive observables the large logarithms outlined above are formed. This is because there will be a miscancellation where virtual interactions can occur in certain phase space regions where real emission cannot.

It is possible for the emitted gluon to be both collinear and soft, which results in a double logarithm. After one additional emission on top of the

[1] In addition to the emission of a gluon from a quark or gluon there is also the interaction where a gluon splits into a quark and anti-quark pair.

hard process the cross section will have the general form

$$\sigma = \sigma^H(1 + \alpha_s(c_1 L^2 + c_2 L + c_3)), \tag{1.9}$$

where L is the large logarithm and c_i are the coefficients. The argument of the logarithms will depend on the choice of the observable. These arguments will in general be proportional to the energies of emitted particles and the angle between the emitter and emitted particles.

1.4.2 Resummation

While a fixed order expansion as outlined above is effective at describing a few additional emissions, it becomes difficult to calculate for higher orders. When a high number of emissions are involved a different approach is required. Instead of calculating each order in turn we can instead look at the general properties that arise in each order. If we take the result which we obtain for the first correction to the hard process and exponentiate it then we will have a first approximation for an all orders result for the cross section. This exponentiation can be justified by considering strong ordering. Each time there is an emission it will be at a scale much less than that of previous emissions. Hence it will be unable to resolve the previous emissions and see the process as it was before any emissions. Roughly speaking, this cross section will be of the form

$$\sigma_r = \sigma^H e^{\alpha_s(c_1 L^2 + c_2 L + c_3)}. \tag{1.10}$$

For observables at an e^+e^- collider this formula is exact. For pp or $p\bar{p}$ interactions additional prefactors are required in order to correctly describe the physical behaviour. This approximation can be improved by matching the coefficients to higher order expansions.

1.4.3 Eikonal limit

Using the full Feynman rules gives the best description of the behaviour of both hard and soft emissions. However the calculational complexity involved with using the full rules can become problematic at higher orders. Most processes will have at most a few hard emissions and many more soft emissions. Hence working in an approximation valid for soft emissions can provide relevant results to many observables. A convenient approach for treating these soft emissions is the eikonal limit.

In the eikonal limit we assume that all of the components of the four-momentum k of an emitted particle are far less than any of the momenta belonging to the particles in the hard scattering p_i,

$$k \ll p_i. \tag{1.11}$$

We now look at what happens when we have a single emission from one of the hard particles. When the emitter is a quark this can be written in terms of Feynman rules as

$$\frac{(\not{p}+\not{k}-m)}{(p+k)^2-m^2}(i\gamma^\mu T^A)u(p) = \frac{iT^A}{(p+k)^2-m^2}(2(p^\mu+k^\mu)-\gamma^\mu\not{k})u(p). \tag{1.12}$$

Now using the above assumption, this becomes

$$\frac{ip^\mu T^A}{k\cdot p}u(p). \tag{1.13}$$

Thus in the eikonal limit the emission amplitude is proportional to the velocity or, in the massless case, the direction of the emitter. This simplified Feynman rule reduces the complexity of the analytical calculation significantly.

1.5 Colour evolution

The colour structure in QCD evolves through the interaction of quarks and gluons. The Feynman rule for the interaction between quarks and gluons is

$$-ig_s\gamma^\alpha T^A_{ik}. \tag{1.14}$$

The Feynman rule has two separate structures. The first structure is the kinematic or Lorentz structure which is contained in the γ^α. This term is irrelevant for considerations of colour evolution. Therefore the only new part of this Feynman rule relative to QED is the SU(3) triplet colour generator T^A_{ik}.

We will first consider the case where the gluon connects two different quark lines. The colour factor in this case is

$$T^A_{ik}T^A_{jl}. \tag{1.15}$$

From the properties of the SU(N_c) generators we can rewrite this expression in terms of Kronecker deltas by using the Fierz identity,

$$T^A_{ik}T^A_{jl} = \tfrac{1}{2}\delta_{il}\delta_{jk} - \frac{1}{2N_c}\delta_{ik}\delta_{jl}. \tag{1.16}$$

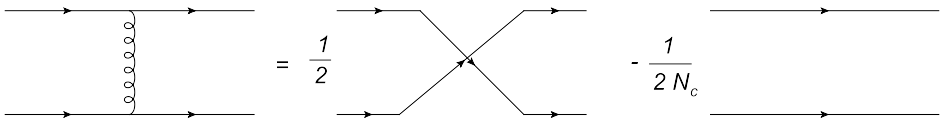

Figure 2: The colour line decomposition of an exchanged gluon.

Each Kronecker delta represents the flow of colour along a quark line. This is shown in Figure 2. By connecting a gluon between two colour lines we generate two new sets of colour lines with different relative weights. Since colour is not an observable quantity, we must sum and average over all possible sets of colour lines when calculating observables.

The extension of the above process to gluon-gluon interactions is slightly more complicated. The two types of gluon-gluon interactions are three and four gluon vertices. The colour structure of the Feynman rule for the triple gluon vertex is

$$if^{ABC} = 2Tr[T^A T^B T^C] - 2Tr[T^B T^A T^C], \tag{1.17}$$

where A, B and C are the colour indices of the gluons. For the triple gluon interaction the complexity arises because there are two unique ways of connecting the quark lines generated by three gluons. The second equality states that the colour structure of a triple gluon vertex can be decomposed into two quark traces. Each of these two quark traces will have a unique colour flow. The colour line representation of the anti-symmetric three gluon coupling is shown in Figure 3.

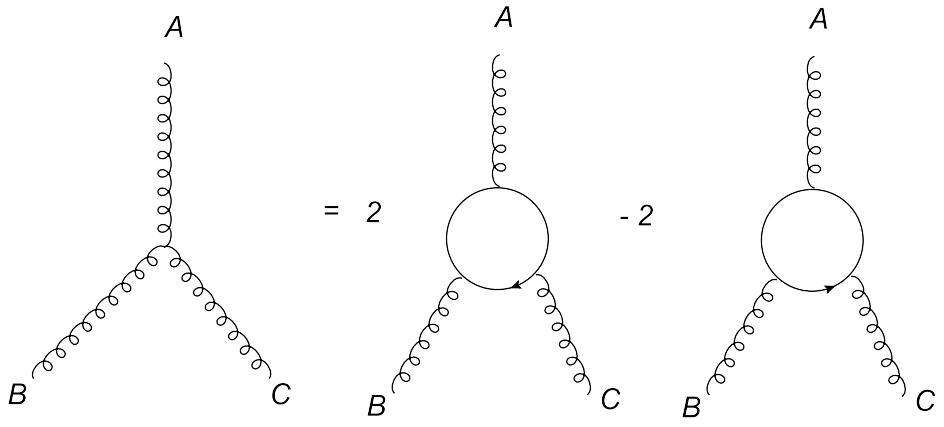

Figure 3: The colour line representation of the anti-symmetric three gluon coupling.

The four gluon interaction is more complicated than the three gluon interaction. This is because it is not possible to separate the colour structure from the Lorentz structure. The four gluon vertex has the

Feynman rule

$$
\begin{aligned}
&-ig_s^2 f^{XAC} f^{XBD} (g^{\alpha\beta} g^{\gamma\delta} - g^{\alpha\delta} g^{\beta\gamma}) \\
&-ig_s^2 f^{XAD} f^{XBC} (g^{\alpha\beta} g^{\gamma\delta} - g^{\alpha\gamma} g^{\beta\delta}) \\
&-ig_s^2 f^{XAB} f^{XCD} (g^{\alpha\gamma} g^{\beta\delta} - g^{\alpha\delta} g^{\beta\gamma}),
\end{aligned} \tag{1.18}
$$

where A, B, C and D are the colour indices of the external gluons, X is a colour index which is summed over, α, β, γ and δ are the Lorentz indices of the external gluons and $g^{\mu\nu}$ is the spacetime metric. It is not possible to separate the kinematics from the colour structures in this case. However, each of the three separate colour structures can be interpreted as a product of two triple gluon vertices, which can be represented by one of the three tree level Feynman diagrams for gluon-gluon scattering.

Each of these colour structures therefore has four possible colour flows and hence each four gluon vertex contains twelve possible colour flows. However it turns out that each colour flow appears in two structures. Hence there are only six unique colour flows.

1.5.1 Leading colour

Often dealing with the full colour structure of QCD is unwieldy, especially when trying to implement such structure in a numerical algorithm. To remove some of the complexities one can use the approximation that the number of colours is large and retain only the leading terms in colour. We will refer to this as the leading colour or large N_c approximation. From the above equations, we can approximate that

$$
T^A_{ik} T^A_{jl} \sim \tfrac{1}{2} \delta_{il} \delta_{jk}. \tag{1.19}
$$

However, one needs to be careful to consider what truly corresponds to the leading colour limit. If we naively remove all terms we believe are

sub-leading then it is possible to remove leading terms as well. A simple example of this is the exchange of colour between two colour lines. The colour factor for this is

$$T^A_{ii} T^A_{jj} = 0. \tag{1.20}$$

since the colour generators are traceless. However if we use the leading colour expression for the colour generators we obtain

$$(T^A_{ii} T^A_{jj})_{N_c \to \infty} = \tfrac{1}{2} N_c. \tag{1.21}$$

A better approach is to let N_c be a variable and perform any calculations using the full colour structure. Once the calculation is complete we can let the number of colours approach infinity. Any result obtained can then be implemented in a numerical algorithm with the guarantee that we are correctly treating the leading colour limit.

1.6 Colour states and evolution

In this section we consider a convenient formalism for expressing the colour structure of a hard process and how it changes under the subsequent evolution. Any matrix element can be expanded in terms of colour basis states C_i as

$$M = A_i C_i, \tag{1.22}$$

where A_i are the Lorentz structures which form the coefficients of the colour basis states. The colour basis states C_i contain different possible configurations of colour lines connecting the external particles. For a set of external particles with equivalent colour representations $\{j, k, l, m\}$ we can write any of the C_i in terms of Kronecker deltas as

$$C_i = \omega^i_1 \delta_{jk} \delta_{lm} + \omega^i_2 \delta_{jl} \delta_{km} + \omega^i_3 \delta_{jm} \delta_{kl}, \tag{1.23}$$

where the ω_j^i factors are functions of N_c and depend on the choice of basis. In order to preserve the direction of colour lines some of the factors ω_j^i must be zero if all of the hard partons are quarks or anti-quarks. Even if the colour representations of the external particles differs, for instance if we have both quarks and gluons, it is possible to construct a basis in a similar form to that above.

Any observable quantity, such as the cross section, will be proportional to the matrix element squared,

$$MM^\dagger = A_i A_j^\dagger (C_i \cdot C_j^\dagger) = Tr[HS], \tag{1.24}$$

where we have defined the hard Lorentz matrix as

$$H_{ij} = A_i A_j^\dagger, \tag{1.25}$$

and the colour metric as

$$S_{ij} = C_i \cdot C_j^\dagger. \tag{1.26}$$

The colour metric contains a sum (average) over final (initial) colour lines as indicated by the inner product. The explicit value of an element S_{ij} is obtained by combining the two states C_i and C_j into a full diagram and calculating the colour factor.

The advantage of the above formalism is that evolving the colour becomes simple. In the eikonal limit we assume that any emission does not change the kinematics of the hard process. Hence any emission of a gluon will only change the colour of the hard process. The emission will also not change the dimensionality of the colour space. The colour structure is purely contained in the colour metric S. The colour metric S_{ij} is calculated by tracing over the combination of the relevant colour states. The hard Lorentz matrix H_{ij} is obtained by calculating the spin traces that occur as prefactors to the colour states. The formalism above

covers only the initial colour structure of the process. The change to the colour structure after an unresolved exchange can be written as

$$\bar{S} = -\xi(\Gamma^\dagger S + S\Gamma), \tag{1.27}$$

where Γ is a colour evolution matrix, often referred to as the anomalous dimension, and ξ is defined as

$$\xi = \int \frac{dt}{t}\frac{2\alpha(t)}{\pi}, \tag{1.28}$$

where $\alpha(t)$ is the strong coupling and t is the scale involved, generally taken to be transverse momentum. The specific form of the anomalous dimension depends on the process in question. After the sum over an arbitrary number of emissions tending towards infinity the colour structure becomes

$$\tilde{S} = e^{-\xi\Gamma^\dagger} S e^{-\xi\Gamma}, \tag{1.29}$$

where Γ is the colour evolution due the emission of a single gluon.

1.7 Colour emission and exchanges

While the kinematics of real emissions and virtual colour exchanges are different, their colour structures are in fact the same. This is a necessity in order to get the cancellations that remove the singularities from the calculations. The colour structure resulting from a possible colour exchange from a hard process is shown in Figure 4. The upper diagram shows the kinematical cut where the gluon is put on-shell. This corresponds to the real emission diagram. The lower diagram shows the kinematical cut where the gluon is left virtual. This corresponds to the virtual exchange of colour. In both cases the colour factor for the diagram is the same. It is independent of the kinematical cut.

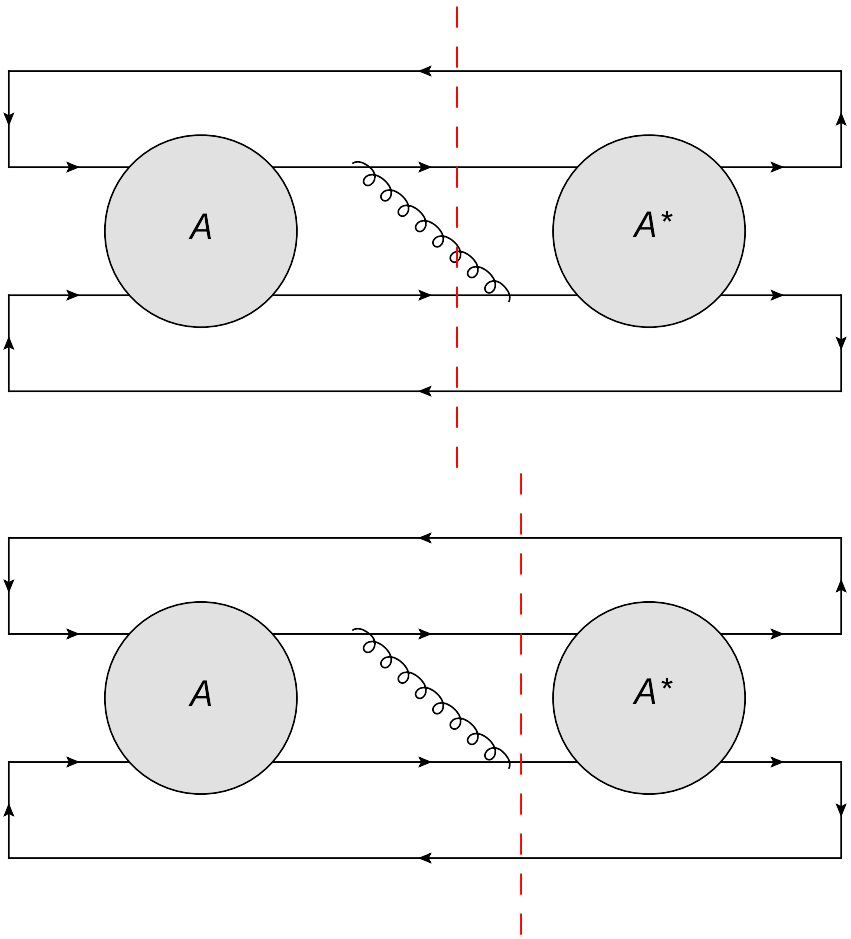

Figure 4: Kinematical cuts for one gluon exchange.

1.8 Colour bases

In order to calculate an observable using QCD one needs to fix the colour basis states. At the end of the calculation the observable quantity will be independant of the basis chosen. Thus we can choose the basis that is most convenient for the specific calculation. In the next two sections we will consider two bases which are convenient for the analytical and numerical approaches respectively.

1.8.1 Orthonormal basis

When one is performing a calculation it is often advantageous to choose a basis with orthonormal states. This way any inner product between two different basis states will be zero, often significantly reducing the

computation required. In terms of the above formalism, this basis is defined by the condition

$$S_{ij} = \delta_{ij}. \tag{1.30}$$

The disadvantage to these bases is that often the states themselves are complicated. While the actual computation of the relevant matrices can be implemented numerically, when one wants to study the leading colour limit the amount of work needed is fairly significant, owing to the complicated basis states.

An example of an orthonormal basis is that for $qg \to qg$. Defining the incoming colour lines as $\{i, B\}$ and the outgoing colour lines as $\{k, D\}$ we can write the basis states e_i for this as

$$e_1 = \delta_{ik}\delta^{BD}, \tag{1.31}$$

$$e_2 = \sqrt{2}if^{BDF}T^F_{ki}, \tag{1.32}$$

$$e_3 = \frac{\sqrt{2}N_c}{\sqrt{(N_c^2-4)}}d^{BDF}T^F_{ki}. \tag{1.33}$$

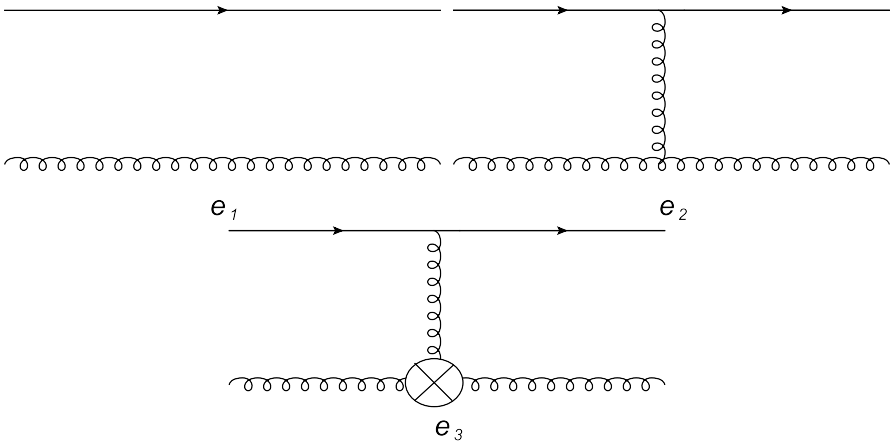

Figure 5: The three orthogonal basis states for $qg \to qg$. The normalization factors are omitted.

A diagrammatic representation of these basis states is given in Figure 5. The prefactors guarantee that the states are normalized to unity. The

third term here is proportional to the colour structure of the symmetric gluon coupling. This coupling does not occur in tree-level perturbation theory, here it is just used as a convenient way of writing

$$d^{ABC} = 2Tr[T^A T^B T^C] + 2Tr[T^B T^A T^C]. \tag{1.34}$$

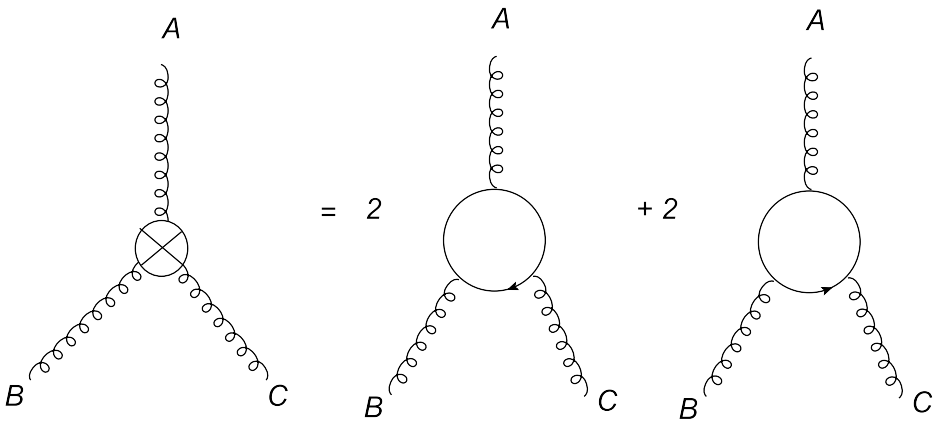

Figure 6: The colour line representation of the symmetric three gluon coupling.

The colour line representation of the symmetric three gluon coupling is shown in Figure 6. It is also used to project out specific colour states when combined with the anti-symmetric gluon coupling. The first colour structure is the singlet exchange. This structure does not occur in any leading order QCD hard process, but is reachable by the exchange of colour lines in the subsequent evolution. The other two basis states contain multiple colour structures.

1.8.2 Colour flow basis

We define the colour flow basis as that which is spanned by the minimal set of simplest possible colour structures. The details of this basis are described in Refs. [1, 2]. This is the basis used in MADGRAPH [3]. Each element of the relevant matrices is simple to calculate, but the basis is

no longer orthogonal and therefore the metric is no longer equivalent to the Kronecker delta. This adds complications when calculating the other quantities, such as the anomalous dimension. However, this is offset by the fact that the simple colour states are often related to each other by symmetries, reducing the number of required computations.

The colour flow basis is more suited to numerical implementation than the orthogonal basis. This is because when we are running a simulation we are considering one fixed colour structure being evolved, rather than averaging over all possible colour states as in the analytical calculation. The colour flow basis also has advantages when performing the analytical calculation. Many of the basis states can be related by interchanging the various partons, reducing the number of elements of the various matrices that need to be calculated.

An example of a colour flow basis is that for $qg \to qg$. Defining the incoming colour lines as $\{i, B\}$ and the outgoing colour lines as $\{k, D\}$ we can write the basis states C_i for this as

$$C_1 = \delta_{ik}\delta^{BD}, \tag{1.35}$$

$$C_2 = \frac{2N_c}{\sqrt{N_c^2 - 1}} T^B_{im} T^D_{mk}, \tag{1.36}$$

$$C_3 = \frac{2N_c}{\sqrt{N_c^2 - 1}} T^D_{im} T^B_{mk}. \tag{1.37}$$

A diagrammatic representation of these basis states is given in Figure 7. The first colour structure is the singlet exchange, as in the orthonormal basis. The other two basis states are the remaining unique colour assignments possible with two quarks and two gluons. The soft metric $S^{CF}_{qg\to qg}$

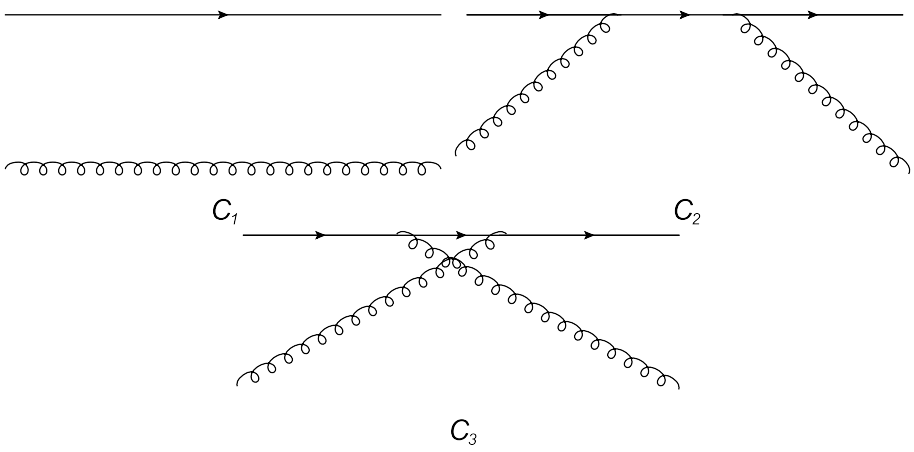

Figure 7: The three colour flow basis states for $qg \to qg$. The normalization factors are omitted.

for this basis is

$$S^{CF}_{qg\to qg} = \begin{pmatrix} 1 & \frac{1}{\sqrt{(N_c^2-1)}} & \frac{1}{\sqrt{(N_c^2-1)}} \\ \frac{1}{\sqrt{(N_c^2-1)}} & 1 & \frac{1}{(N_c^2-1)} \\ \frac{1}{\sqrt{(N_c^2-1)}} & \frac{1}{(N_c^2-1)} & 1 \end{pmatrix}. \tag{1.38}$$

1.8.3 Relating the bases

The two bases can be related by a transformation matrix. Since we have complicated colour structures the transformations are not trivial rotations, except in the leading colour limit. The transformation between the colour flow and orthonormal basis is

$$\begin{pmatrix} C_1 \\ C_2 \\ C_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{\sqrt{N_c^2-1}} & \frac{N_c}{\sqrt{2(N_c^2-1)}} & \sqrt{\frac{N_c^2-4}{2(N_c^2-1)}} \\ \frac{1}{\sqrt{N_c^2-1}} & -\frac{N_c}{\sqrt{2(N_c^2-1)}} & \sqrt{\frac{N_c^2-4}{2(N_c^2-1)}} \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix}. \tag{1.39}$$

1.8.4 Special Case: $gg \to gg$

While we have outlined above that one can obtain the colour states of a basis by finding all possible ways of connecting external colour lines,

there is one special case that we need to consider. For $gg \to gg$ there are two possible vertices that we can use to construct states. We can either directly connect the gluons or we can connect the gluons to a quark loop. Since the resulting states are simpler we will only consider the quark loop method here, but it is possible to construct a set of basis states with the three gluon symmetric and anti-symmetric vertices [4].

There are three unique ways of connecting the gluons without any vertices. These are the s, t and u channel singlets. In addition to these there are six ways of connecting the four external gluons with a quark loop that differ from the singlets. Three of these states are identical to another state apart from the direction of the internal quark loop. From a symmetry perspective these states must be averaged over and hence the true states are a sum of the two with opposite loop directions. Thus only six states are needed to describe $gg \to gg$ at leading order. In the colour flow basis these states are

$$C_1 = \delta^{AC}\delta^{BD}, \tag{1.40}$$

$$C_2 = \delta^{AD}\delta^{BC}, \tag{1.41}$$

$$C_3 = \delta^{AB}\delta^{CD}, \tag{1.42}$$

$$C_4 = DTr[T^A T^C T^D T^B + T^B T^D T^C T^A], \tag{1.43}$$

$$C_5 = DTr[T^A T^C T^B T^D + T^D T^B T^C T^A], \tag{1.44}$$

$$C_6 = DTr[T^A T^D T^C T^B + T^B T^C T^D T^A], \tag{1.45}$$

where D is a normalisation factor given by,

$$D = \sqrt{\frac{8N_c^2(N_c^2-1)}{(N_c^4 - 2N_c^2 + 6)}}. \tag{1.46}$$

We have neglected the states similar to C_4 to C_6 but with a minus signs intead of the plus signs. These states do not occur in the hard processes

and in fact decouple from the six states listed above [5]. Hence they do not affect physical observables in any way. It should be noted that in the case of $N_c = 3$ we find that it is possible to express one of these states in terms of the other five [6]. This results in an overcomplete basis if we use six basis states and take the limit. However, it is perfectly acceptable to use six basis states with general N_c, in fact it is required.

1.9 Numerical Simulation of QCD

While the analytical resummation can give a good picture of the physical phenomena, often a numerical approach is required. The most common approach for numerically simulating high energy physics is to use a Monte-Carlo event generator. The most popular event generators in use at the moment are Herwig++ [7–9], PYTHIA6 [10], PYTHIA8 [11] and SHERPA [12]. An event generator consists of several components, each corresponding to a physical model. The first component is a hard process generator, which chooses the hard process of an interaction with a probability given by QCD calculations. The hard process is then evolved by allowing the partons to split. In addition to the primary hard scattering additional lower energy scatters may also be generated. When all processes have been generated and evolved the final coloured partons are then converted in to colourless hadrons by a hadronization mechanism. A schematic representation of the processes within the event generator is given in Figure 8.

In this section we will discuss the parts of the generator with regards to only QCD $2 \to 2$ scattering processes. Many more processes can be simulated in the most popular event generators, but these are assumed to be less affected by any changes in the parton shower evolution.

An in-depth review of event generators is given in Ref. [13].

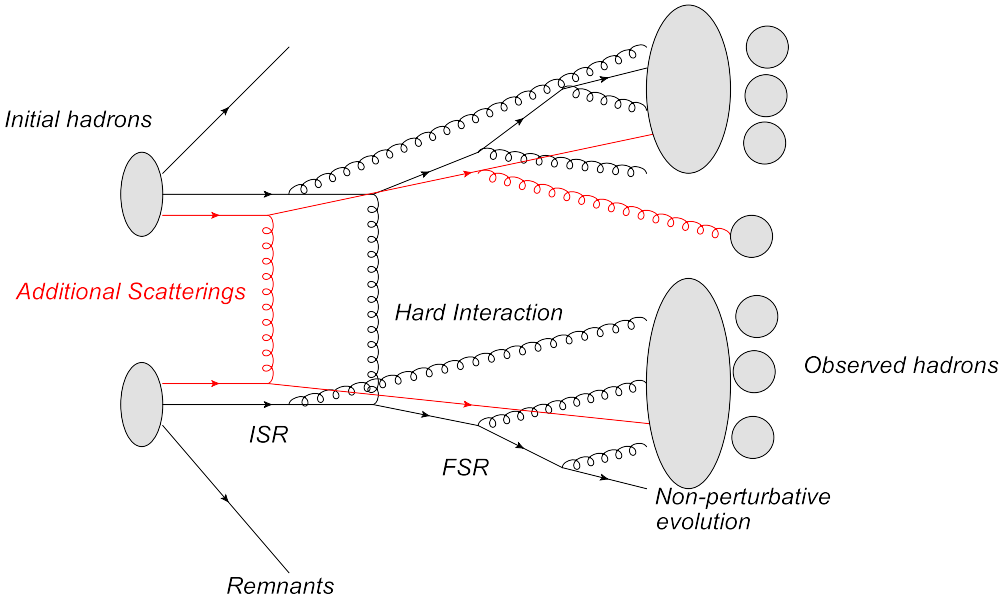

Figure 8: A schematic representation of an event generator.

1.9.1 Hard Process

The hard process is selected by using random numbers with all possible processes weighted by their physical phase space, PDF and kinematic factors. In addition to selecting the flavours of the partons involved in the interaction, the colour structure of the hard process needs to be considered.

The purpose of assigning the colour structures in the hard process is to determine the maximum scales used in the parton shower. While in the analytical approach it is allowed for the amplitude and conjugate to have different colour structures, this is difficult to implement numerically. If we simply removed the non-diagonal colour structures then we would not reproduce the correct cross section for the hard process. Instead, we choose a different approach first used in Ref. [14]. Utilizing the previous formalism, we can define a new hard matrix

$$\bar{H}_{ij} = \delta_{ij}(S_{ii}H_{ii} + \sum_{k\neq i} S_{ik}H_{ik}), \tag{1.47}$$

where indices are not summed over unless explicitly stated. This form of the hard matrix guarantees the condition

$$Tr[\bar{H}] = Tr[HS]. \tag{1.48}$$

The colour structure i is then chosen using the kinematic weight $\bar{H}_{ii}$. While the full colour structure, taking in to account the differences above, is used for the hard process, the evolution of colour without making any further approximations is extremely complicated. The methods used for dealing with some of the terms which are sub-leading in colour will be explained in the Chapter 4. Current implementations do not include sub-leading colour terms, they merely assume that the colour structure is constant and described by colour lines connecting partons which do not change under the consequent evolution in the parton shower. That is, they assume that the anomalous dimension matrix is diagonal in the colour flow basis.

1.9.2 Parton Shower

Once the hard process has been decided, we need to simulate the radiation generated by the scattering of colour lines.

There are three different approaches to the parton shower. The first approach, which we will refer to as the traditional parton shower, considers each of the partons in the hard process to evolve independently. Each of the splittings produces children that are on-shell but have undetermined momentum. At the end of the splittings for a parton the necessary information to determine the momenta of all of the children is obtained. At this point, since all final-state partons produced in the splitting are on-shell it is possible to evolve another of the partons from the hard process. This is the approach used by Herwig++. The second approach, which we will call a dipole shower, considers colour connected dipoles to

split, rather than the individual partons. In this approach one dipole will split in to two dipoles, and each parton involved will be on-shell at all times. This has the advantage over the parton shower that the shower can be stopped at any time and all particles will be physical. This is the approach used by SHERPA. A third approach, which we will refer to as a hybrid shower, behaves like the parton shower, in that it treats individual partons as splitting, but properly sets particles on shell by adding a colour connected recoil particle, like in the dipole approach. This is the approach used by PYTHIA8. In this section we will focus on how the parton and hybrid showers evolve the partons from the hard process.

Evolution is handled by defining a scale t proportional to the virtuality, transverse momentum or some other measure of hardness of a given kinematical setup. The initial scale t_i is determined from the hard process. It is usually related to the virtuality or transverse momentum of colour connected pairs of partons. The parton shower is an evolution in the variable t, from t_i down to a final scale t_f. The final scale t_f is typically chosen to be around 1 GeV, which is where the assumptions used in the parton shower are no longer valid. As the scale is evolved there is a probability of radiation at the given scale t.

The parameters for the next splitting are solved by equating the Sudakov factor $\Delta(t, t_n)$, which describes the probability of non-emission between scales t and t_n, to a random number R_1 in the range [0,1],

$$\Delta(t, t_n) = R_1, \tag{1.49}$$

where $\Delta(t, t_n)$ for final-state splittings is given by

$$\Delta(t, t_n) = \exp\left(- \int_{t_n}^{t} \frac{1}{2\pi} \frac{dt'}{t'} \int_{\epsilon_-}^{\epsilon_+} \alpha_s(z, t') P(z, t') dz \right), \tag{1.50}$$

where $P(z, t')$ is a splitting kernel which contains all possible methods

of splitting the current parton or dipole. The limits of the momentum fraction z, denoted $\epsilon_{\pm}$, in Herwig++ are obtained by solving

$$z(1-z)t^2 = -m_{ij}^2 + \frac{m_i^2}{z} + \frac{m_j^2}{(1-z)} - \frac{p_t^2}{z(1-z)}, \tag{1.51}$$

where m_{ij} is the mass of the splitting parton, m_i and m_j are the masses of the partons produced by the splitting and p_t is the transverse momentum. In principle, depending on the form of the splitting kernel and expression used for α_s, it may be possible to solve this expression analytically. However, the complicated boundaries on z makes such a solution difficult. In practice, we instead define overestimate functions α_s^{over} and $P^{over}(z)$ such that the above expression can be solved. The overestimates used are simplified versions of the DGLAP splitting kernels [15–17]. Using these overestimates we can define

$$I(z) = \int P^{over}(z), \tag{1.52}$$

and therefore obtain the solution for the next scale

$$t_n = tR_1^{\delta}, \tag{1.53}$$

where δ is defined as

$$\delta = \frac{2\pi}{\alpha_s^{over}} \frac{1}{(I(\epsilon_+) - I(\epsilon_-))}. \tag{1.54}$$

To completely describe a splitting we also need two additional parameters. We choose the first to be the energy fraction z, which is the fraction of momentum which the primary child parton gains from the parent. The value of z is obtained by solving

$$\int_{\epsilon_-}^{z} P(z')dz' = R_2 \int_{\epsilon_-}^{\epsilon_+} P(z')dz', \tag{1.55}$$

where R_2 is another random number in the range [0,1]. As above, we use the same overestimate of the splitting kernel in order to analytically rearrange for z

$$z = I^{-1}(R_2 I(\epsilon_+) + (1 - R_2) I(\epsilon_-)), \tag{1.56}$$

where $I^{-1}(x)$ is the inverse function of the integration $I(z)$. In order to recover the correct splitting probabilities we need apply a veto to the splitting. After finding values for t and z we veto the splitting if

$$R_3 > \frac{P(z,t)}{P_{over}(z)}, \tag{1.57}$$

$$R_4 > \frac{\alpha_s(z,t)}{\alpha_s^{over}}, \tag{1.58}$$

where R_3 and R_4 are random numbers in the range [0,1]. The final parameter we need is the azimuthal angle of the transverse momentum generated in the splitting. It is usually chosen with a random number R_5 in the range [0,1], as

$$\phi = 2\pi R_5. \tag{1.59}$$

The actual colour structure of splittings in the shower is determined using the leading colour limit. For $q \to qg$ and $g \to q\bar{q}$ the colour assignment from the parent to the children particles is unique. For $g \to gg$ there are two possibilities which are shown in Figure 9. Depending on the scales involved one of the two possibilities may be the physically correct choice, or it is possible that neither is preferred and a random choice is more appropriate. This is discussed in more detail in Chapter 2.

The parton shower continues evolving partons until they reach a minimum scale. One choice of minimum scale is to require that the transverse momentum generated in a splitting is greater than 1 GeV. Since it is difficult to experimentally distinguish two partons with very

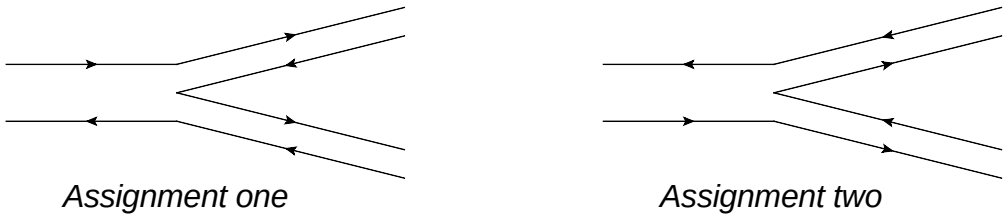

Figure 9: The two possible colour line assignments in $g \to gg$.

small relative transverse momentum there is no need to continue with the shower at that point.

1.9.3 Multiple Interactions

As the energies and luminosities of the beams in a collider increase it is possible for interactions to involve multiple scatterings. There are two kinds of multiple interactions that can occur. The first is multiple scatterings involving different protons. Essentially many separate events happen at the same beam crossing. In general very few of the multiple scatterings will be interesting, so removing the less important events from a more important event is often required. If we had a perfect detector then it would be possible to separate out all of these events. Since this is a purely experimental phenomena it is not included in event generators.

The second type of multiple scattering is one where two beam particles interact multiple times with each other. It is unlikely that a high scale interaction will be accompanied by additional scatterings with comparable scales. Unlike the previous case these are part of the underlying theory and not due to the limitations of detectors. In order to correctly simulate the phenomenology these additional scatterings must be included in the event generator [18–21].

1.9.4 Hadronization

The parton shower generates the high multiplicity events we observe in the experimental data. However, we have never observed a free quark or gluon in any experiment. Instead we have only observed colourless hadrons. There must be some mechanism that transforms these partons in to the colourless hadrons observed in particle detectors. Since these transformations involve the low energy particles at the end of the shower we cannot calculate the mechanism perturbatively. Instead we must rely on physically inspired models to describe the transition. There are currently two different models used, one based on colour connected clusters [22] and one based on colour strings [23].

At the end of the parton shower there will be a mixture of quarks, anti-quarks and gluons in the final state. In the cluster model any remaining gluons are first transformed into quark and anti-quark pairs. All of the colour connected quark and anti-quark pairs form an intermediate object, which is referred to as a cluster. All of the connections between the colour lines were formed in the parton shower. These colour lines determine the properties of the clusters and the distribution of the final-state hadrons.

In recent versions of Herwig++ a colour reconnection model has been added to the hadronization process [24]. In this model the colour lines assigned in the parton shower can be changed when certain criteria are met.

1.10 Conclusion

While there are still aspects of QCD which require further study we understand enough to make high precision predictions for experimentally observable quantities. These predictions are made using both theoretical calculations and numerical simulations. However, the current algorithms used in the numerical simulation of QCD lack much of the complexity

of colour contained within the analytical approach. In order for the numerical approach to remain precise enough to accurately predict future experimental results we must begin to implement this missing behaviour. In the next three chapters we will outline how changing the algorithms used and performing a tuning may improve the predictive power of our event generators.

References

[1] F. Maltoni, K. Paul, T. Stelzer, and S. Willenbrock, "Color flow decomposition of QCD amplitudes," *Phys.Rev.* **D67** (2003) 014026, arXiv:hep-ph/0209271 [hep-ph].

[2] M. Sjödahl, "Color structure for soft gluon resummation: A General recipe," *JHEP* **0909** (2009) 087, arXiv:0906.1121 [hep-ph].

[3] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, "MadGraph 5 : Going Beyond," *JHEP* **1106** (2011) 128, arXiv:1106.0522 [hep-ph].

[4] Y. Dokshitzer and G. Marchesini, "Soft gluons at large angles in hadron collisions," *JHEP* **0601** (2006) 007, arXiv:hep-ph/0509078 [hep-ph].

[5] N. Kidonakis, G. Oderda, and G. F. Sterman, "Evolution of color exchange in QCD hard scattering," *Nucl.Phys.* **B531** (1998) 365–402, arXiv:hep-ph/9803241 [hep-ph].

[6] A. MacFarlane, A. Sudbery, and P. Weisz, "On Gell-Mann's lambda-matrices, d- and f-tensors, octets, and parametrizations of SU(3)," *Commun.Math.Phys.* **11** (1968) 77–90.

[7] M. Bähr *et al.*, "Herwig++ Physics and Manual," *Eur. Phys. J.* **C58** (2008) 639–707, arXiv:0803.0883 [hep-ph].

[8] S. Gieseke *et al.*, "Herwig++ 2.5 Release Note," arXiv:1102.1672 [hep-ph].

[9] K. Arnold, L. d'Errico, S. Gieseke, D. Grellscheid, K. Hamilton, *et al.*, "Herwig++ 2.6 Release Note," arXiv:1205.4902 [hep-ph].

[10] T. Sjöstrand, S. Mrenna, and P. Z. Skands, "PYTHIA 6.4 Physics and Manual," *JHEP* **0605** (2006) 026, arXiv:hep-ph/0603175 [hep-ph].

[11] T. Sjöstrand, S. Mrenna, and P. Z. Skands, "A Brief Introduction to PYTHIA 8.1," *Comput.Phys.Commun.* **178** (2008) 852–867, arXiv:0710.3820 [hep-ph].

[12] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, *et al.*, "Event generation with SHERPA 1.1," *JHEP* **0902** (2009) 007, arXiv:0811.4622 [hep-ph].

[13] A. Buckley, J. Butterworth, S. Gieseke, D. Grellscheid, S. Hoche, *et al.*, "General-purpose event generators for LHC physics," *Phys.Rept.* **504** (2011) 145–233, arXiv:1101.2599 [hep-ph].

[14] K. Odagiri, "Color connection structure of supersymmetric QCD (2 $\rightarrow$ 2) processes," *JHEP* **9810** (1998) 006, arXiv:hep-ph/9806531 [hep-ph].

[15] V. N. Gribov and L. N. Lipatov, "Deep inelastic e p scattering in perturbation theory," *Sov. J. Nucl. Phys.* **15** (1972) 438–450.

[16] G. Altarelli and G. Parisi, "Asymptotic Freedom in Parton Language," *Nucl. Phys.* **B126** (1977) 298.

[17] Y. L. Dokshitzer, "Calculation of the Structure Functions for Deep Inelastic Scattering and e^+e^- Annihilation by Perturbation Theory

in Quantum Chromodynamics," *Sov. Phys. JETP* **46** (1977) 641–653.

[18] M. Bahr, S. Gieseke, and M. H. Seymour, "Simulation of multiple partonic interactions in Herwig++," *JHEP* **0807** (2008) 076, arXiv:0803.3633 [hep-ph].

[19] T. Sjöstrand and P. Z. Skands, "Multiple interactions and the structure of beam remnants," *JHEP* **0403** (2004) 053, arXiv:hep-ph/0402078 [hep-ph].

[20] T. Sjöstrand and P. Z. Skands, "Transverse-momentum-ordered showers and interleaved multiple interactions," *Eur.Phys.J.* **C39** (2005) 129–154, arXiv:hep-ph/0408302 [hep-ph].

[21] R. Corke and T. Sjöstrand, "Multiparton Interactions and Rescattering," *JHEP* **1001** (2010) 035, arXiv:0911.1909 [hep-ph].

[22] B. Webber, "A QCD Model for Jet Fragmentation Including Soft Gluon Interference," *Nucl.Phys.* **B238** (1984) 492.

[23] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjöstrand, "Parton Fragmentation and String Dynamics," *Phys.Rept.* **97** (1983) 31–145.

[24] S. Gieseke, C. Rohr, and A. Siodmok, "Colour reconnections in Herwig++," *Eur.Phys.J.* **C72** (2012) 2225, arXiv:1206.0041 [hep-ph].

2 Jet Vetoes and Herwig++

2.1 Introduction

In the previous chapter we discussed the colour structure of QCD and how it impacts phenomenology. In this chapter we will consider one particular implementation of QCD in a theoretical model. The model we will consider is the parton shower described in Chapter 1.

In this chapter we will look at an experimental observable known as gaps-between-jets. This observable is important to investigate due to its dependence on wide-angle radiation. Thus predictions for this observable from an event generator will be highly dependent on the physics implemented in the parton shower.

In addition to general interest in the observable we should also look into the odd behaviour found in Ref. [1]. In this paper it was shown that the results for the parton shower in Herwig++ [2–4] for this observable did not match the leading colour analytical calculation. Instead, it seemed to match the full colour analytical calculation, which includes effects that are not included in the parton shower. This odd behaviour clearly needs to be investigated.

Furthermore, the colour structure of this observable is the same as more complex processes, such as the production of a Higgs boson plus two jets. By studying this observable it is possible that we can reduce the QCD background encountered when looking for such processes [5] or, at least, the uncertainty on its prediction.

The discussions in the chapter are mostly taken from Ref. [6].

2.2 Gaps Between Jets - Analytical Approach

To study the behaviour of the event generator we should consider a specific observable which is heavily dependent on the implementation of QCD in

the parton shower. One such observable is where we set a requirement that there is a limited amount of radiation in a specific region of phase space. This observable is commonly referred to as "gaps-between-jets" or dijet production with a rapidity gap. This observable has been studied by many groups [1,7–9] and measured at HERA [10–12], the Tevatron [13–15] and the LHC [16]. The process is shown in Figure 10. For this observable the restricted region is the phase space between either the two hardest jets or the most forward and backward jets. In this chapter we will only consider the case where the phase space is bound by the two hardest jets. We will refer to this region of phase space as the "gap".

The observable is defined in terms of three quantities. The first is the size of the rapidity gap Y. As the rapidity gap increases there is a larger amount of phase space between the two hard jets. As a result of this there is a greater possibility of emissions between the two hard jets. The second quantity is the hard scale Q. We choose to define this as the average transverse momentum of the two hardest jets. Increasing the value of this scale will allow more energy to be radiated in to the restricted region. The final parameter is the veto scale Q_0. An event with a gap is one which has no jets with transverse momentum above the veto scale in the region bounded by the two hardest jets.

We can use the colour basis formalism outlined in Chapter 1. The resummed partonic cross section has the form

$$\sigma_{Gap} = \mathrm{Tr}[He^{-\xi\Gamma^\dagger}Se^{-\xi\Gamma}], \tag{2.1}$$

where S is the colour metric, H is the hard matrix, Γ is the anomalous dimension and ξ is defined by

$$\xi = \int_{Q_0}^{Q} \frac{dk_t}{k_t}\frac{2\alpha_s(k_t)}{\pi} \simeq \frac{2\alpha_s}{\pi}\log(\tfrac{Q}{Q_0}), \tag{2.2}$$

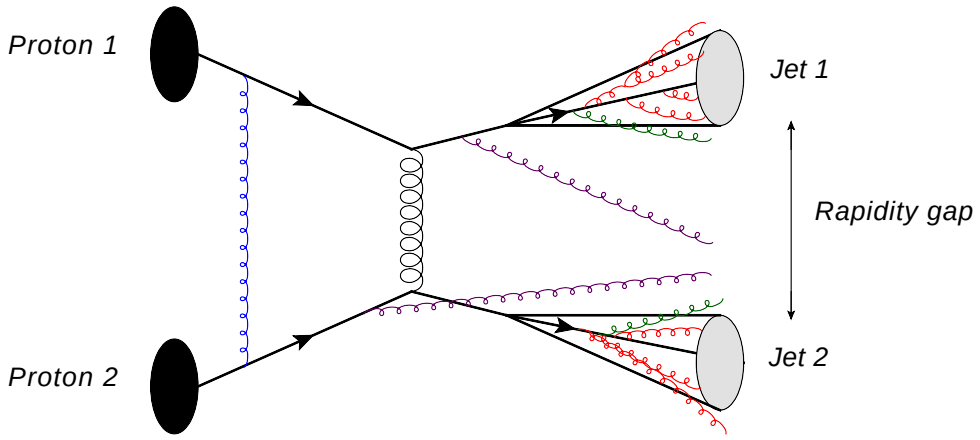

Figure 10: Dijet production with a rapidity gap. The black lines represent the partonic interaction and those with arrows can be quarks, antiquarks or gluons. The blue gluon represents a virtual exchange which occurs in the analytical calculation but not in the parton shower. The purple gluons represent wide-angle radiation that enters the rapidity gap. The green gluons represent small-angle radiation that is just able to reach the rapidity gap. The red gluons are small-angle radiation that does not enter the gap.

where we have assumed that the strong coupling α_s is constant for illustrative purposes. This expression can be modified to include the running coupling. In this equation we see that limiting the amount of radiation allowed in the gap has generated a large logarithm of the ratio of the hard scale of the boundary jets and the veto scale. This is a result of the miscancellation between the virtual correction, which is allowed to exchange gluons of any energy in the gap, and the real emission, which is limited to energies below the veto scale.

The actual hard matrices and soft metric depend on the partonic process. The total resummed partonic cross section will be the sum of the cross sections for each of the individual QCD $2 \to 2$ scattering processes. For this observable the anomalous dimension can be written

in the form [17]

$$\Gamma = \tfrac{1}{2} Y \mathrm{t}_t^2 + i\pi \mathrm{t}_1 \cdot \mathrm{t}_2 + \tfrac{1}{4}\rho(\mathrm{t}_3^2 + \mathrm{t}_4^2). \tag{2.3}$$

Here t_t^2 is the colour exchange across the gap and t_i are the colour generators for the four different hard partons i. The first term is due to wide-angle radiation emitted into the gap. As the gap size is increased there is a greater possibility of radiation and therefore a reduction in the overall gap cross section. The second term is from virtual exchanges of colour between the initial or final state partons. These colour exchanges cannot be observed but do change the overall colour structure of the event. The third term is a small effect where small-angle radiation is able to make it in to the gap. The magnitude of this term depends on the definition of jets used. The factor ρ is given by

$$\rho = \log\left(\frac{\sinh(Y+R)}{\sinh(R)}\right) - Y, \tag{2.4}$$

where R is the jet radius. The elements of the anomalous dimension for any arbitrary colour basis can be calculated using

$$\Gamma_{ij} = \sum_k S_{ik}^{-1} \langle C_k \mid \Gamma \mid C_j \rangle. \tag{2.5}$$

An example diagram used to calculate the elements is shown in Figure 11. The example is for an on-diagonal element of $qq \to qq$. This diagram only corresponds to the element of the anomalous dimension in the orthonormal basis. It must be multiplied by the inverse of the soft metric in order to get the true anomalous dimension in a general basis.

The cross section outlined above is correct only at the partonic level. To generate the true cross section it must be convoluted with the parton distribution functions (PDFs). However it is possible to choose the

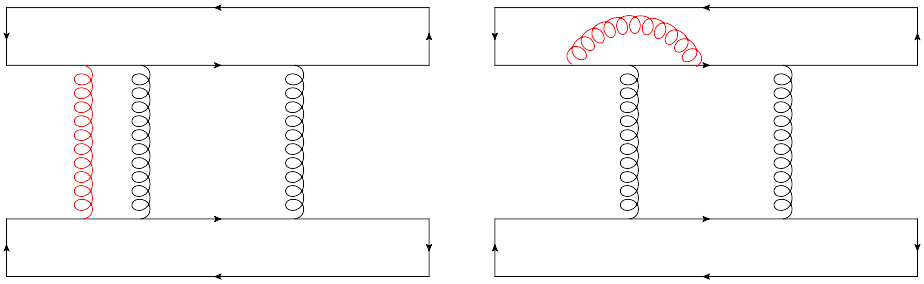

Figure 11: Left: The colour diagram for the $t_1 \cdot t_2$ term. Right: The colour diagram for the $t_1 \cdot t_3$ term, which is contained in t_t^2. The exchanged gluon is shown in red.

variables in such a way that the PDFs can be shuffled into a luminosity pre-factor multiplying the cross section. The full resummed gap cross section can then be written as

$$\sigma_H(Q, Q_0, Y) = \sum_i \mathcal{L}_i(Q, Y)\sigma_{Gap,i}(Q, Q_0, Y). \tag{2.6}$$

where the sum is over all possible QCD $2 \to 2$ interactions and $\mathcal{L}_i(Q, Y)$ is a factor dependent on the PDFs involved in process i.

To calculate the values of H, S and Γ we choose to work in the colour flow basis and let N_c be a variable. This way it will be possible to work in both the $N_c \to 3$ and $N_c \to \infty$ limits using the same calculation.

2.3 Parton shower differences

The algorithms used in the parton shower have been discussed in Chapter 1. In this section we will discuss how the implementation of colour in the parton shower differs from that of the analytical calculation. The specific parton shower that we will consider is that of Herwig++. However, our discussion here also applies to the parton shower in PYTHIA 6 [18].

The parton shower is formulated mostly in terms of the leading colour approximation. There is one sub-leading colour effect which is included,

which is the colour charge of the quark being C_F rather than $\frac{1}{2}N_c$. No other sub-leading effects are included in the parton shower and these will be discussed in Chapter 4.

The parton shower is an evolution in a scale from some initial scale down to a smaller cut-off scale. The scale in Herwig++ is related to the angle between, and energies of, a colour connected pair of partons. More details on this are given in Ref. [19]. The initial scale of the parton shower is determined by the colour structure of the hard process. The final scale of the parton shower is usually fixed to a value of the order 1 GeV.

For quarks and anti-quarks there is only one colour or anti-colour line respectively and therefore only one possible choice of scale. For gluons there are two possible scales since the two colour lines are always connected to different partons in leading colour QCD $2 \to 2$ interactions [1]. Since the two scales may be quite different it is important that there is a method for determining which one should be used.

In the default implementation one of the two colour lines is chosen at random. Each colour line has a fifty percent chance of being chosen. In QCD $2 \to 2$ interactions in the leading colour limit at least one of the colour lines attached to the gluon will cross the gap and therefore be able to emit wide-angle radiation. Each gluon can therefore be connected to either two wide-angle lines or a wide-angle line and a small-angle line. Examples of the two cases are shown in Figure 12. In addition to these two colour structures shown for $gg \to gg$ there is also a third one which has a similar number of lines crossing the gap as the right diagram.

In the first case it is not very important which of the two lines is picked since the scales will generally be similar for small-angle scatterings

[1] In the case of the colour singlets the colour lines of the gluon are always attached to one other gluon in the event. In this case both scales would be the same. It is possible that neither line will cross the gap which will result in minimal radiation in to the gap. The effects of these singlets will be discussed in Chapter 4.

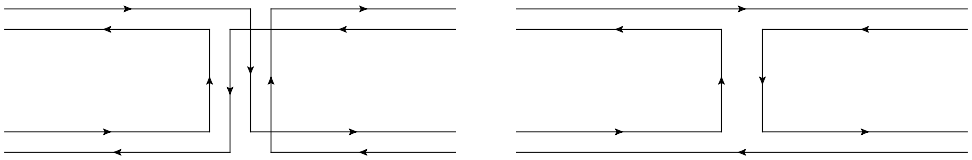

Figure 12: Left: Each gluon is connected to two wide-angle colour lines. Right: Each gluon is connected to one wide-angle colour line and one small-angle colour line. Here we have assumed that the scattering is small-angle.

and therefore generate similar behaviour. In the second case the choice of line is more important. This case is shown in Figure 13. If the wide-angle line is chosen then radiation has the possibility to be produced in the gap with a colour factor N_c. This is twice as much radiation per gluon as we would expect. If the small-angle line is chosen instead then no wide-angle radiation will be produced at all. After averaging over the two or four gluons in the $2 \to 2$ process this model should on average produce the right amount of radiation at wide-angles. However, there is still the possibility that we will end up with more or less radiation than is predicted from analytical calculations.

When the parton shower generates a splitting of the form $g \to gg$ there is ambiguity in how the colour lines are distributed. The parton shower evolves partons rather than colour lines. Since the gluon has a colour and an anti-colour line we must determine which of the two to attach radiation to. The leading colour structure for the splitting is shown in Figure 14. It is not clear from this diagram which of the two lines has emitted a new gluon. The default implementation chooses the emitter at random from one of the two colour lines. Each of the two colour lines has a fifty percent chance of being chosen.

It is important to note that the random choice of initial hard colour lines does not occur in dipoles or dipole-inspired parton showers such

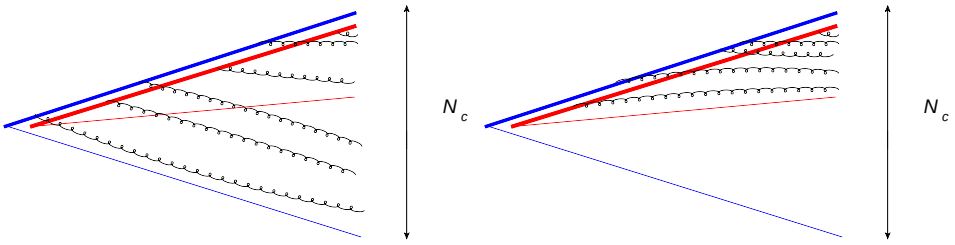

Figure 13: Left: The partner with the largest scale is chosen, which results in wide-angle radiation being produced with a colour factor N_c. Right: The partner with the smaller scale is chosen, which results in only small-angle radiation produced with a colour factor N_c. The blue and red lines represent the colour connections to the higher scale and lower scale partners respectively. The bold lines are those of the gluon and the fainter lines are those of the two partners.

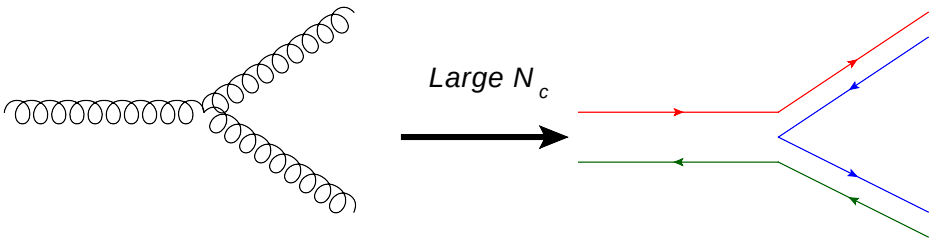

Figure 14: Colour line representation of $g \to gg$ splitting in the leading colour limit.

as SHERPA [20], MATCHBOX [21], VINCIA [22] and PYTHIA 8 [23]. Hence it is very important to study these effects so that Herwig++ is able to compete with these event generators.

2.4 Analytical Modification

2.4.1 Analytical "parton shower" model

It is important to understand how the different assumptions within the parton shower will have an effect on experimental observables. From a qualitative point of view we could guess that, depending on which colour

line is chosen to set the scale, there could be more or less radiation per event than intended. In order to have a quantitative understanding of the relative sizes of these effects we need a more mathematical approach.

To this end we will modify our analytical approach in such a way that it includes the effects outlined in the previous section. We will consider two analytical models. Both analytical models will treat the colour structure of the hard process in the same way as Herwig++. The hard matrices of the analytical calculation will be modified in the way outlined in Chapter 1. Both analytical models will work in the leading colour limit apart from the quark colour charge which will remain C_F.

The difference between the two analytical models will be in the resummation, specifically in the form of the anomalous dimension. Since the colour structure of the parton shower is fixed by the hard process we will assume that the off-diagonal elements of the anomalous dimension vanish, as is true in the leading colour limit. All that remains is to determine what values the on diagonal terms should take. In the first model the colour factors in the anomalous dimension from the analytical model will be retained, but will be treated in the leading colour limit. In the second model the colour factors in the anomalous dimension will be derived from the behaviour of the parton shower.

The parton shower works in the leading colour limit. In this limit quarks can be regarded as a colour line and gluons can be regarded as a colour and an anti-colour line. The anomalous dimension has contributions arising from the radiation of gluons from the quarks and gluons in the hard process. The anomalous dimension can be written as

$$\Gamma = \Gamma_\rho + \sum_q \Gamma_q + \sum_g \Gamma_g \longleftrightarrow e^{-\Gamma} = e^{-\Gamma_\rho} e^{-\sum_q \Gamma_q} e^{-\sum_g \Gamma_g}. \tag{2.7}$$

Here we have assumed that the anomalous dimension is diagonal, which is true when using the colour flow basis in the leading colour limit. The

factor Γ_ρ is a suppression factor to the total cross section due to collinear radiation just making it in to the gap. The effect from this factor is usually much smaller than the suppression due to wide-angle radiation in the other terms. The factors Γ_q and Γ_g represent the contributions from quarks and gluons respectively. For every colour line crossing the gap a factor

$$\Gamma_i = \tfrac{1}{4} C_i Y \tag{2.8}$$

is added to the anomalous dimension, where C_i is equal to C_F if the colour line is attributed to a quark or $\frac{1}{2}N_c$ if the colour line is attributed to a gluon[2].

There are two ways in which we can implement the 50-50 choice of hard colour lines for a gluon in the hard process. The first method is to double the size of the anomalous dimension for every gluon in the hard process and add every possible combination of the anomalous dimension. For $gg \to gg$ this results in 96 possible anomalous dimension values. A more elegant method is to make the approximation

$$e^{-2\Gamma_g} \simeq \tfrac{1}{2}(1 + e^{-4\Gamma_g}), \tag{2.9}$$

which is true to first order in Γ_g. This way we see that there will be no emission from gluons half the time and twice as much wide-angle radiation the other fifty percent of the time. For the process $gg \to gg$ there are four gluons. If we let the rapidity gap become extremely large this results in a cross section proportional to the factor

$$e^{-2\Gamma_\rho}(\tfrac{1}{2} + \tfrac{1}{2}e^{-4\Gamma_g})^4 \to \frac{1}{16}e^{-\xi C_A \rho} \simeq \frac{1}{16}\left(\frac{Q}{Q_0}\right)^{-\frac{2\alpha_s C_A \rho}{\pi}}, \tag{2.10}$$

[2]We consider the colour lines separately for each parton. The actual radiation from a colour line crossing the gap will be the sum of the colour factors of the two partons that is connected by said line.

where in the last equality we have assumed that the strong coupling is constant. Clearly this factor is non-zero, even for arbitrarily large gaps. This is the behaviour that a singlet exchange would generate. Since singlet exchanges are not included in the parton shower, this is unwanted behaviour. We will refer to this behaviour as a "quasi-singlet". The correct analytical formalism has a resummed cross section proportional to

$$e^{-2\Gamma_\rho} e^{-8\Gamma_g} \sim \sum_i \left(\frac{Q}{Q_0}\right)^{-\frac{2\alpha_s C_A(\rho + A_i Y)}{\pi}}. \tag{2.11}$$

Here A_i is always greater than zero as some lines must cross the gap in the leading colour limit. As a result the resummed cross section vanishes much more quickly at high rapidity gaps, which is expected.

2.4.2 Partonic Fractions

To visualize the different behaviour in the parton shower we will now look at how the different processes contribute to the resummed cross-section. We will define a quantity called the partonic fraction ω_i as the ratio of a resummed gap cross section for one process i, σ_i^r, to the total resummed gap cross section as

$$\omega_i = \frac{\sigma_i^r}{\sum_i \sigma_i^r}. \tag{2.12}$$

The partonic fractions for pp interactions with beam energies of 7 TeV, $Y = 1$ and $Q = 500$ GeV are shown in Figure 15. The upper plot shows the behaviour of the analytical model in the leading colour limit and the lower plot shows the analytical model with the modifications from the parton shower. The processes with the highest cross sections are those with poles in the t and u channel. The processes with only s channel poles will only have minor contributions. Since the rapidity gap is small we should see similar behaviour for both the analytical and parton shower models. We can see that for high Q_0 this is indeed the case, but for lower

Q_0 there are some deviations between the two models. We can see that in the analytical model the $gg \to gg$ partonic fraction falls off slightly faster than in the parton shower model. The difference at this rapidity gap size is quite minimal. To look into this deviation further, we will now consider the case of a more extreme rapidity gap.

The partonic fractions for pp interactions with beam energies of 7 TeV, $Y = 5$ and $Q = 500$ GeV are shown in Figure 16. It is clear that at the greater Q_0 values the two approaches appear to be roughly equal.

We would expect that as the Q_0 cutoff is lowered the processes with more colour exchanged across the gap should have a faster reduction in the cross section than those with less colour exchanged. This is because the dominant term in the anomalous dimension is proportional to $t_t^2 Y$. For processes with with only quarks or anti-quarks only two colour lines can cross the gap. Similarly, processes which are a mix of gluons and quarks can only have two colour lines crossing the gap.

Interactions with four gluons in the hard process have a possibility that four colour lines can cross the gap. This possibility is shown on the left of Figure 12. For small-angle scatterings all four colour lines will cross the gap. However, this term is also suppressed kinematically at small angles as it is proportional to a u-channel pole. In the opposite case, where we have scatterings at almost 180 degrees, only two colour lines will cross the gap for this term. Hence this configuration where four colour lines cross the gap is only important when the soft scale reaches unrealistically small values.

But this is not what happens in the analytical model of the parton shower. Instead the processes with more gluons tend to have a slower reduction in cross section than those with only quarks. This is due to the configurations where gluons can choose colour partners on the same side of the gap, reducing the amount of radiation produced.

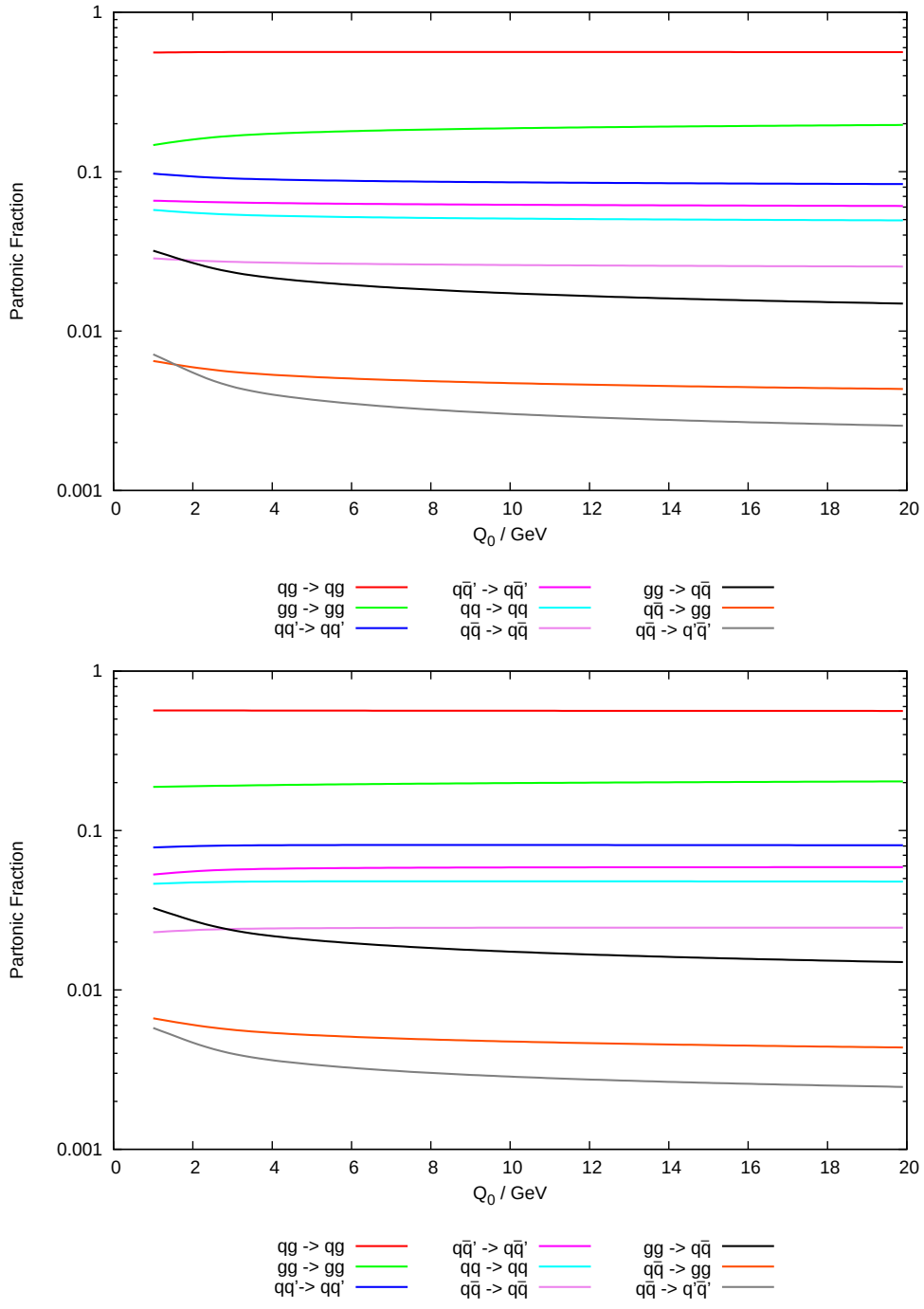

Figure 15: Partonic fractions of the resummed cross section as a function of veto scale for Y=1, Q=500 GeV. The upper plot is the analytical model in the leading colour limit. The lower plot is the analytical model of Herwig++, which contains the changes to colour evolution outlined previously. The PDF sets used are MSTW08lo [24].

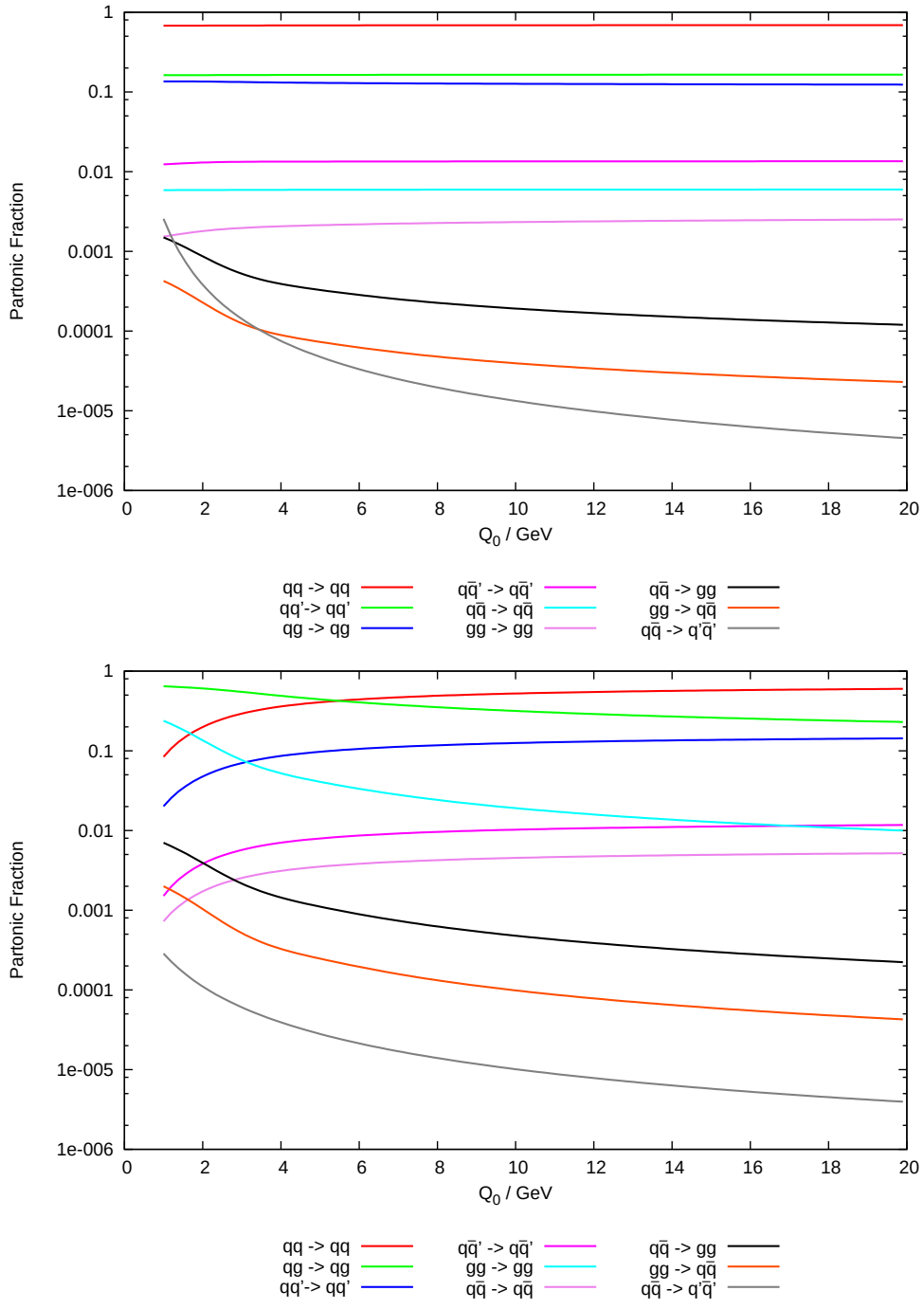

Figure 16: Partonic fractions of the resummed cross section as a function of veto scale for Y=5, Q=500 GeV. The upper plot is the analytical model in the leading colour limit. The lower plot is the analytical model of Herwig++, which contains the changes to colour evolution outlined previously. The PDF sets used are MSTW08lo.

2.4.3 Energy-momentum conservation

One difference between the analytical approach and the parton shower that we have not yet considered is energy and momentum conservation. In the traditional parton shower energy and momentum are not conserved exactly at each splitting. Instead, once all of the children of a hard parton have been evolved down to the minimum scale the energy and momentum are recalculated so that each parton is on shell.

In the analytical approach we are using the eikonal model. As such, we have assumed that each emission from the hard partons carries no momentum at all. This is clearly different than the parton shower. To include the full effects of energy and momentum conservation would be difficult as we could no longer use the eikonal formalism.

Rather than modify the hard matrices, we consider the additional energy required in order to have an extra emission inside the gap. This would modify the energy fractions x_i used in the PDFs. The most general form would be to allow an emission of an additional gluon with any transverse momentum k_t less than the gap limit. The new energy fractions $\overline{x}_i$ would be of the form

$$\overline{x}_{1,2} = x_{1,2} + \frac{k_t e^{\pm(y' + \frac{\overline{y}}{2})}}{\sqrt{s}}. \tag{2.13}$$

This form makes the analytical calculation much more difficult as transverse momentum is no longer split from the PDFs. Instead we consider the maximum effect of one unresolved emission, which is taking $k_t = Q_0$. The new energy fractions become

$$\overline{x}_{1,2} = x_{1,2} + \frac{Q_0 e^{\pm\frac{Y+\overline{y}}{2}}}{\sqrt{s}}. \tag{2.14}$$

It is important to note that as long as Q is much greater than Q_0 this shift in energy would not be large,

$$\frac{\overline{x}_{1,2}}{x_{1,2}} = 1 + \mathcal{O}\left(\frac{Q_0}{Q}\right). \tag{2.15}$$

This method of imposing energy and momentum conservation is only a basic approximation. It can be an overestimate of the effects of a single emission or an underestimate of the effects of multiple emissions inside the gap. A similar approach was also used in Ref. [25].

The effects of imposing the basic energy and momentum conservation modification outlined above are shown in Figure 17. The gradient is much steeper at low Q_0 when energy and momentum conservation is included. This is because when only low energy emissions are allowed inside the gap the factors $\overline{x}$ are barely changed, but when high energy emissions are allowed the PDFs give a major suppression.

2.5 Parton shower modification

Having shown that the current parton shower has unrealistic behaviour for low values of the soft scale, we must now make a modification to our algorithms. To implement better behaviour we propose the following changes.

The first change is that the highest scale colour line will always be chosen to set the scale of the parton shower. This guarantees that there is always a possibility of wide-angle radiation if there are colour lines crossing the gap. The parton shower will then proceed with the evolution, but with half the normal colour factor ($\frac{1}{2}N_c$) since only one colour line can emit. Once the evolution scale is below that of both hard colour lines it is possible for both colour lines to emit and the colour factor is restored to N_c.

In addition to the changes regarding scales and colour factors there is also an additional change. Whenever radiation is emitted it will be attached to the correct lines. For wide-angle emissions only the wide-angle

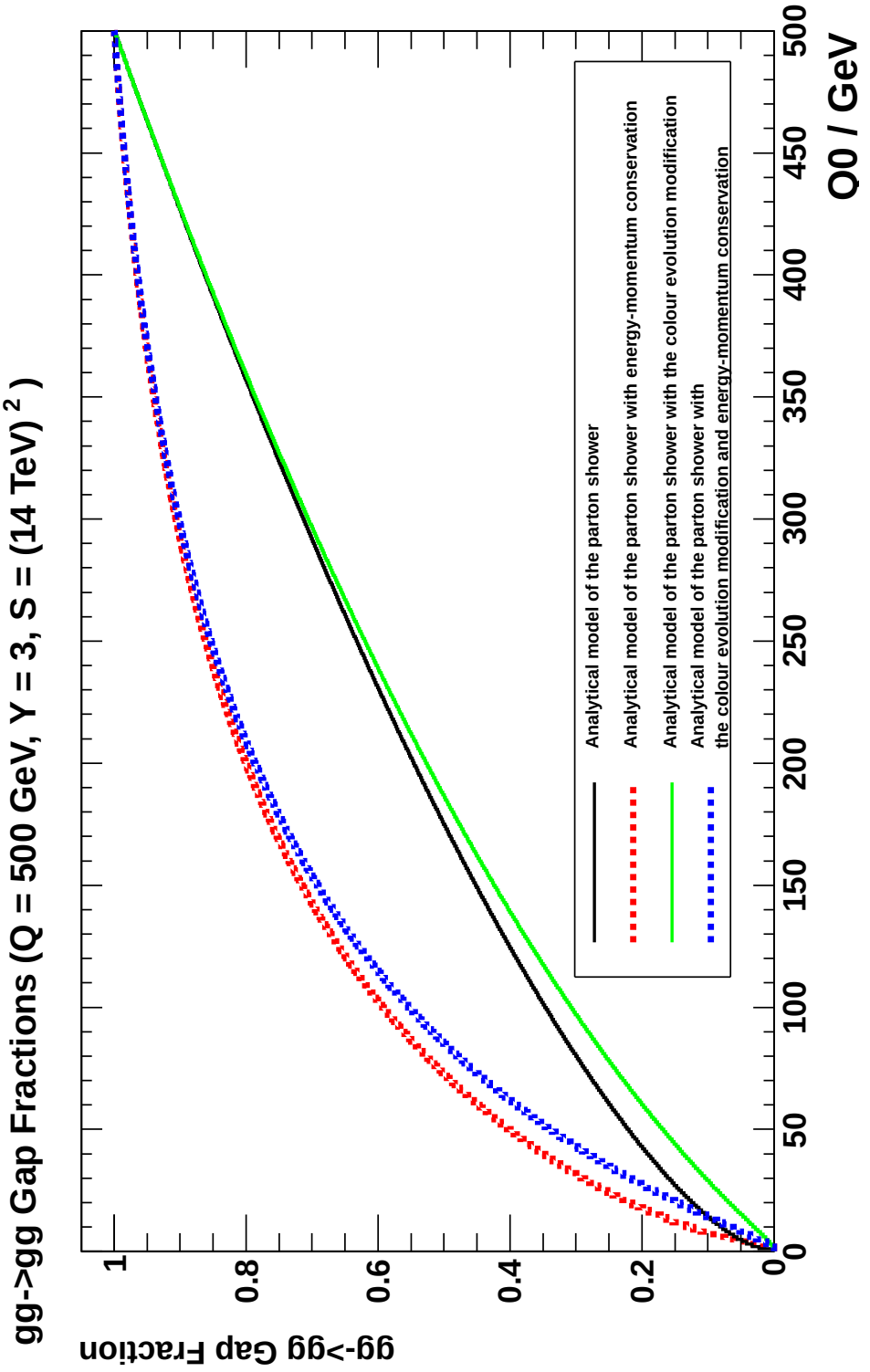

Figure 17: Effects of partial energy-momentum conservation in the analytical calculation on the $gg \to gg$ gap fraction. The implementation of energy-conservation increases the gap fraction for low Q_0, as observed in the transition from black to red and green to blue respectively.

line will be allowed to emit. For small-angle emissions the colour line will be picked from one of the two possibilities with a 50-50 chance. A possible radiation pattern emerging from this new behaviour is shown in Figure 18.

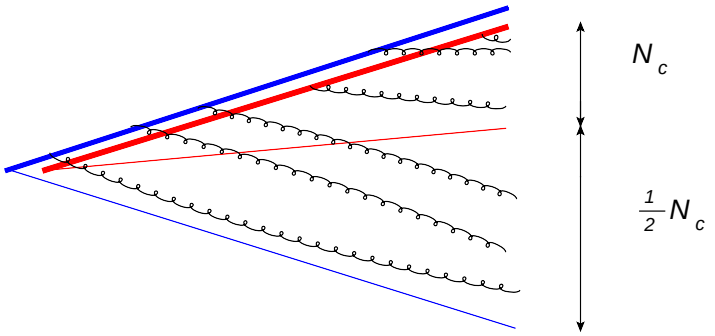

Figure 18: The new radiation pattern for the modified Herwig++ parton shower. The blue and red lines represent colour connections to the higher and lower scale partners respectively. The bold lines are those of the gluon and the fainter lines are those of the two partners.

2.6 Results

Now that we have implemented our changes in the parton shower we can see the impact that it has on phenomenology. We will specifically consider the gaps between jets observable as outlined previously. For these results we use the MSTW08lo PDF set [24] and use SISCONE [26] via FastJet [27] with cone radius R=0.4 and overlap parameter 0.5 for jet finding. This is done in order to make a comparison to the original analysis [1] which highlighted the problems that we have been investigating. Multiple interactions are turned off for all of these results. The event generator used is Herwig++ 2.5.2 with modifications to include the new colour evolution.

We will first check that the behaviour of our implementation is correct. In order to do this we will need to run Herwig++ without the hadroniza-

tion model. This is required since the analytical approach only represents the parton shower and does not include hadronization corrections. We will look at $gg \to gg$ gap fractions since these are the processes which are most impacted by the changes we have made. The results are shown in Figure 19. These results include the energy and momentum conservation modifications to the analytical calculation.

We can see that there is a clear difference between the numerical results and the analytical calculation. However, there are some differences between the parton shower and the analytical approach that we have still not included. Instead of trying to include additional effects, we can instead look at ratios of these curves instead to see if we have implemented the correct changes. The results are in Figures 20 and 21. The ratios show that we do indeed have the correct implementation.

Now that we have established that our implementation is correct we can look at the effects of our modification to hadronization. This is shown in Figure 22. We see that the gap fraction increases at lower Q_0 values in both cases when hadronization is included. However, with the modification we see a greater increase relative to the unmodified version. This is due to wide-angle radiation no longer being attached to small-angle lines, resulting in less hadrons being produced in the wide-angle region.

Having studied the effects of the two modifications that we have made we can now show our final results for the gap cross section. This was our original motivator for the study. This is shown in Figure 23. We include a run from Herwig++ 2.3.0 in order to verify the behaviour previously shown in Ref. [1].

It is clear that Herwig++ 2.5.2 fits better without our modification to the colour evolution than with it. The unmodified result almost matches the expected Forshaw-Keates-Marzani result [1], while the modified result is consistently lower. In order to get the correct behaviour from our event

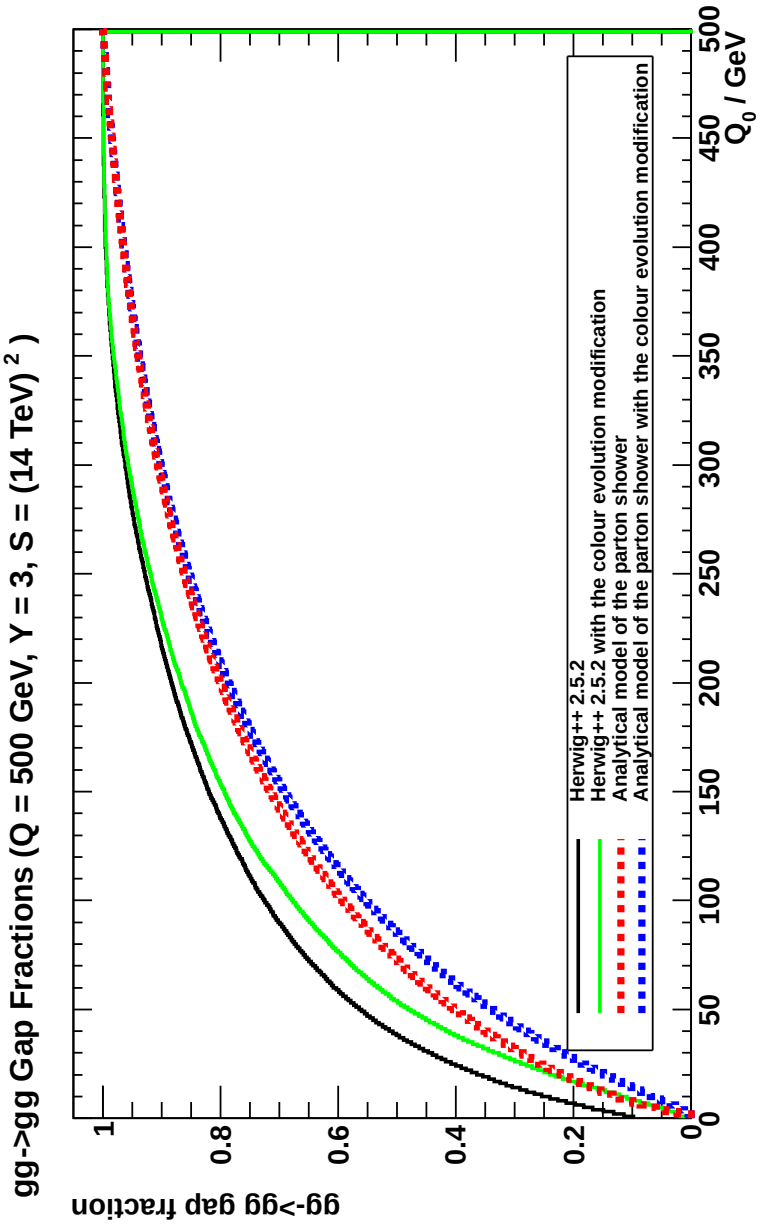

Figure 19: The gap fraction as a function of Q_0. The red and black curves are the numerical results with and without the modifications to the colour evolution. The green and blue curves are the analytical results with and without the modifications to the colour evolution, both including the modifications to account for energy-momentum conservation described in Section 2.4.3.

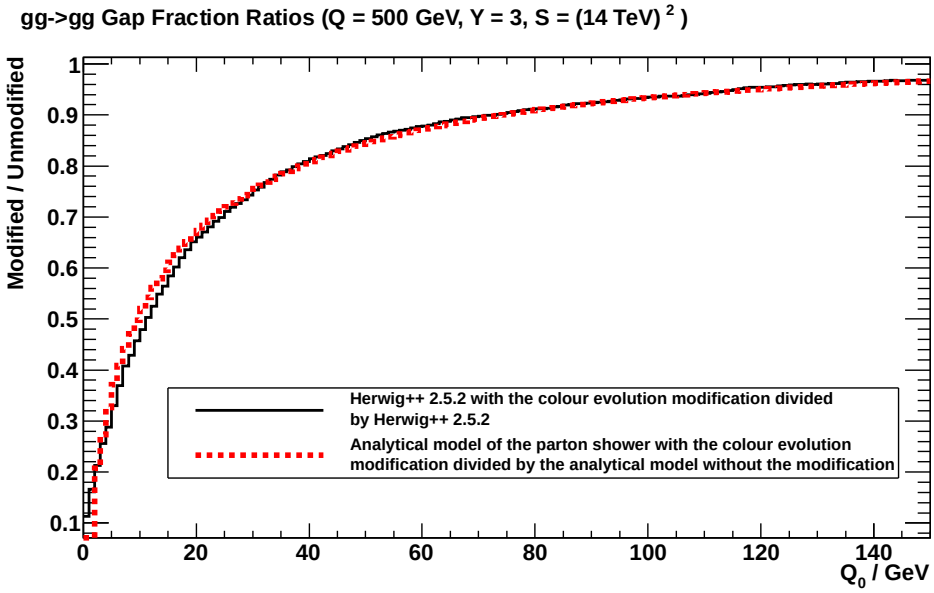

Figure 20: The ratio of the colour evolution modified $gg \to gg$ gap fractions to those without modifications for Herwig++ 2.5.2 (black) and the analytical model of the parton shower (red).

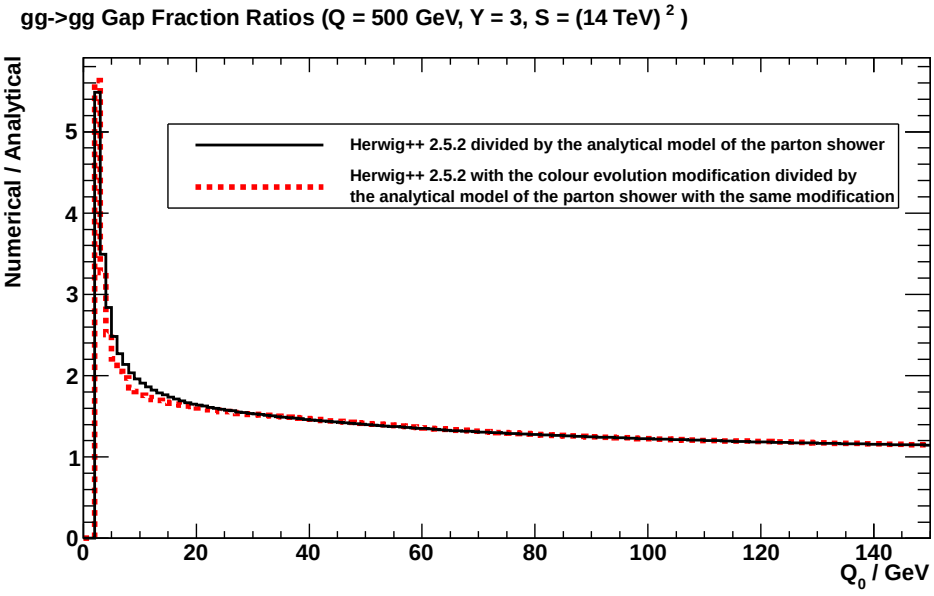

Figure 21: The ratio of Herwig++ 2.5.2 to the analytical model of the parton shower with the colour evolution modification (black) and without (red).

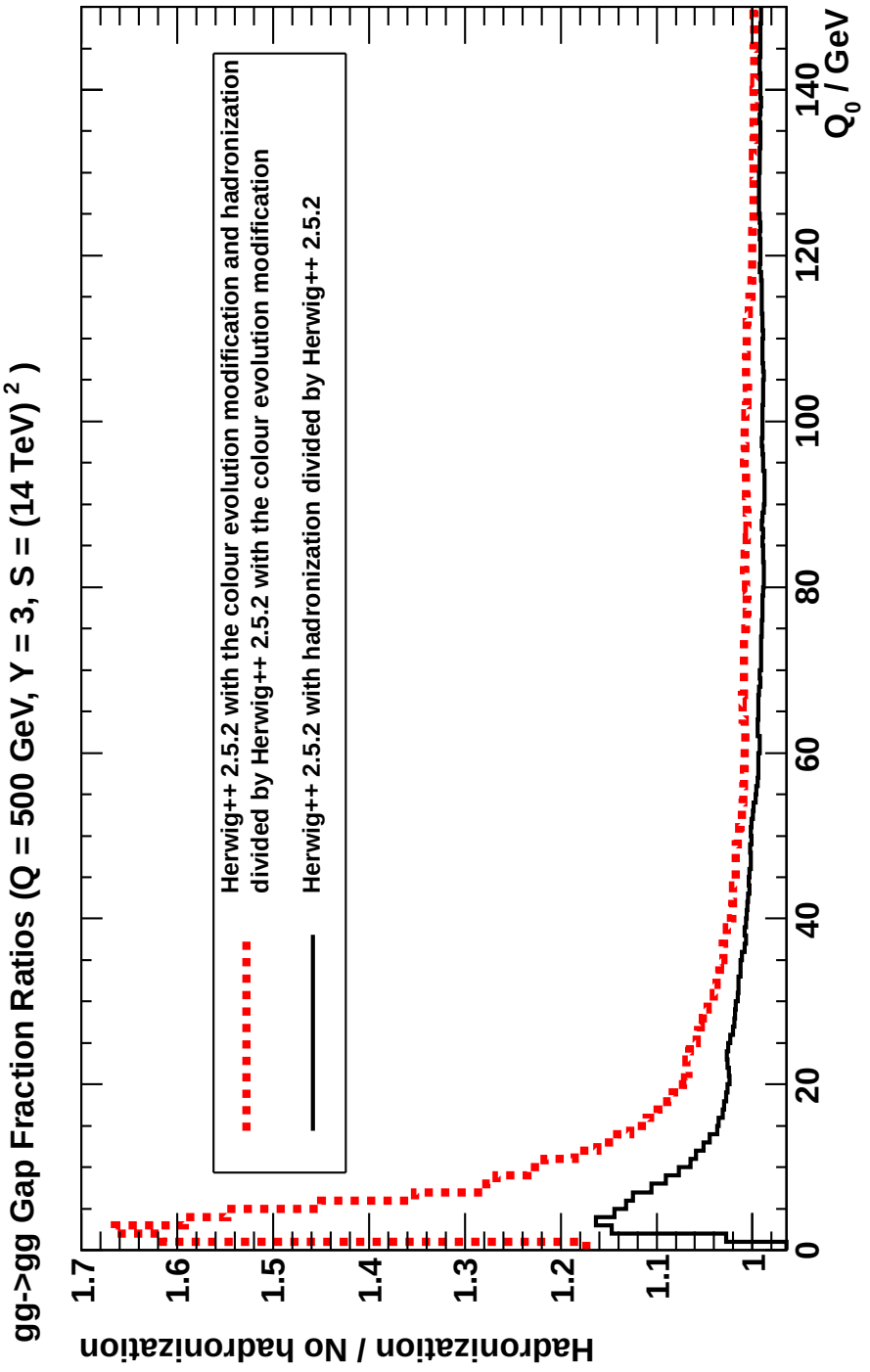

Figure 22: The ratio of the $gg \to gg$ gap fraction generated by Herwig++ with hadronization to that generated by Herwig++ without hadronization.

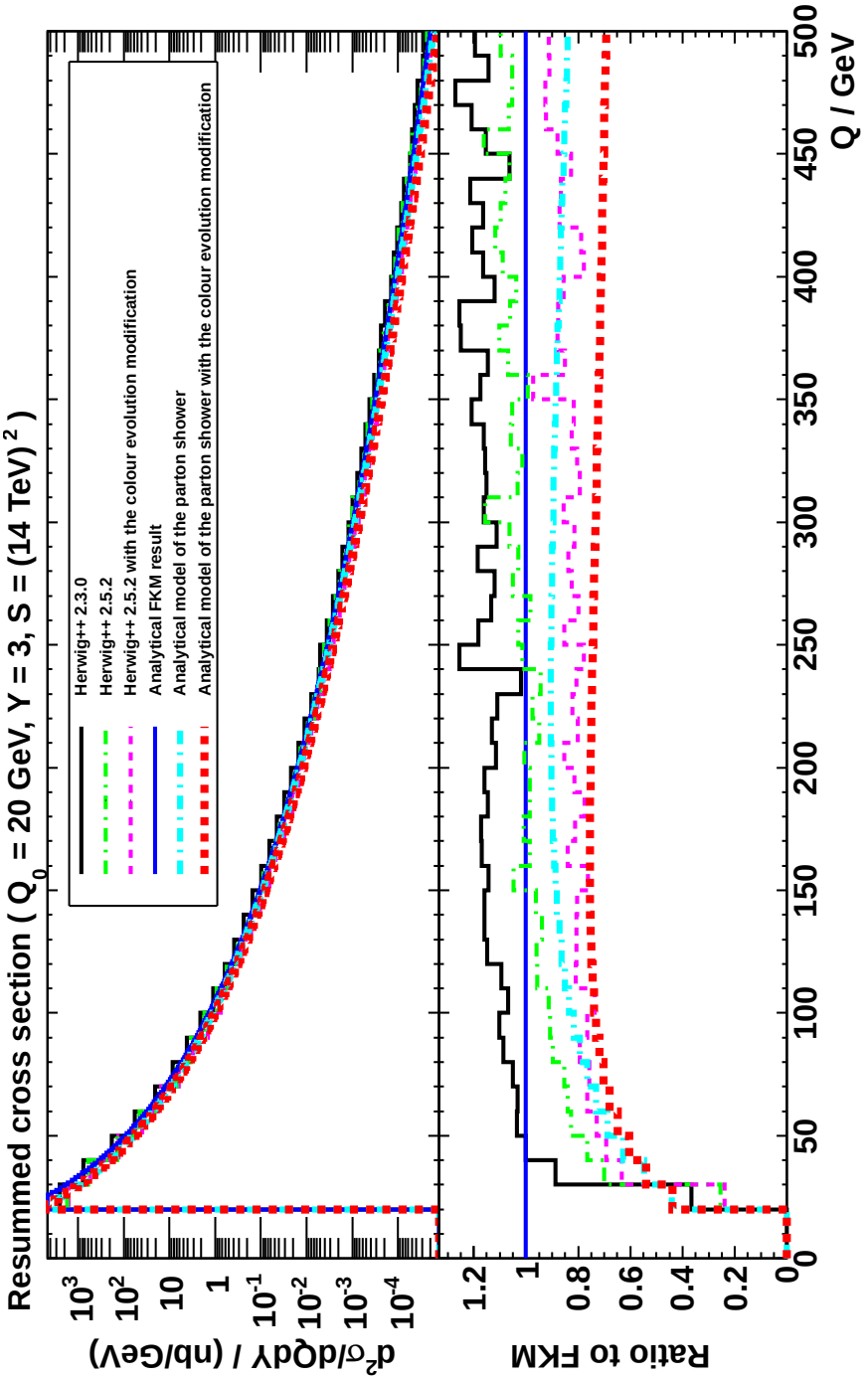

Figure 23: The gap cross section for all possible QCD $2 \to 2$ processes as a function of the hard scale Q. The FKM result is the analytical result without the imaginary terms from Ref. [1].

generator we may now have to retune the parameters of some of our models to incorporate the effects of our new evolution method.

2.7 Conclusion

It is important that the algorithms in our event generators implement the correct physics. While this is not always possible due to the complexity of the algorithms involved we should ensure that the implementation of any physics that we do include is done correctly.

We have shown in this chapter that, by changing how the scales are chosen at the start of the parton shower, there are potential changes to the results of observables predicted by the event generator. These changes will hopefully allow our event generator to give a better description of QCD. These changes are included as an option as part of the newest version of Herwig++ [4].

Having implemented our change to the parton shower it is important that we now attempt to retune our models in the event generator to see if it is possible to get a better representation of physics. We will do this in Chapter 3.

References

[1] J. Forshaw, J. Keates, and S. Marzani, "Jet vetoing at the LHC," *JHEP* **07** (2009) 023, `arXiv:0905.1350 [hep-ph]`.

[2] M. Bähr *et al.*, "Herwig++ Physics and Manual," *Eur. Phys. J.* **C58** (2008) 639–707, `arXiv:0803.0883 [hep-ph]`.

[3] S. Gieseke *et al.*, "Herwig++ 2.5 Release Note," `arXiv:1102.1672 [hep-ph]`.

[4] K. Arnold, L. d'Errico, S. Gieseke, D. Grellscheid, A. Schofield, *et al.*, "Herwig++ 2.6 Release Note," arXiv:1205.4902 [hep-ph].

[5] B. E. Cox, J. R. Forshaw, and A. D. Pilkington, "Extracting Higgs boson couplings using a jet veto," *Phys.Lett.* **B696** (2011) 87–91, arXiv:1006.0986 [hep-ph].

[6] A. Schofield and M. H. Seymour, "Jet vetoing and Herwig++," *JHEP* **1201** (2012) 078, arXiv:1103.4811 [hep-ph].

[7] N. Kidonakis, G. Oderda, and G. F. Sterman, "Evolution of color exchange in QCD hard scattering," *Nucl. Phys.* **B531** (1998) 365–402, arXiv:hep-ph/9803241.

[8] C. F. Berger, T. Kucs, and G. F. Sterman, "Energy flow in interjet radiation," *Phys. Rev.* **D65** (2002) 094031, arXiv:hep-ph/0110004.

[9] R. Appleby and M. Seymour, "The Resummation of interjet energy flow for gaps between jets processes at HERA," *JHEP* **0309** (2003) 056, arXiv:hep-ph/0308086 [hep-ph].

[10] **ZEUS Collaboration** Collaboration, M. Derrick *et al.*, "Rapidity gaps between jets in photoproduction at HERA," *Phys.Lett.* **B369** (1996) 55–68, arXiv:hep-ex/9510012 [hep-ex].

[11] **H1 Collaboration** Collaboration, C. Adloff *et al.*, "Energy flow and rapidity gaps between jets in photoproduction at HERA," *Eur.Phys.J.* **C24** (2002) 517–527, arXiv:hep-ex/0203011 [hep-ex].

[12] **ZEUS Collaboration** Collaboration, S. Chekanov *et al.*, "Photoproduction of events with rapidity gaps between jets at HERA," *Eur.Phys.J.* **C50** (2007) 283–297, arXiv:hep-ex/0612008 [hep-ex].

[13] **CDF Collaboration** Collaboration, F. Abe *et al.*, "Observation of rapidity gaps in $\bar{p}p$ collisions at 1.8 TeV," *Phys.Rev.Lett.* **74** (1995) 855–859.

[14] **D0 Collaboration** Collaboration, B. Abbott *et al.*, "Probing hard color-singlet exchange in $p\bar{p}$ collisions at $\sqrt{s} = 630$ GeV and 1800 GeV," *Phys.Lett.* **B440** (1998) 189–202, arXiv:hep-ex/9809016 [hep-ex].

[15] **CDF Collaboration** Collaboration, F. Abe *et al.*, "Dijet production by color - singlet exchange at the Fermilab Tevatron," *Phys.Rev.Lett.* **80** (1998) 1156–1161.

[16] **ATLAS Collaboration** Collaboration, G. Aad *et al.*, "Measurement of dijet production with a veto on additional central jet activity in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector," *JHEP* **1109** (2011) 053, arXiv:1107.1641 [hep-ex].

[17] J. R. Forshaw, A. Kyrieleis, and M. H. Seymour, "Super-leading logarithms in non-global observables in QCD: Colour basis independent calculation," *JHEP* **09** (2008) 128, arXiv:0808.1269 [hep-ph].

[18] T. Sjöstrand, S. Mrenna, and P. Z. Skands, "PYTHIA 6.4 Physics and Manual," *JHEP* **0605** (2006) 026, arXiv:hep-ph/0603175 [hep-ph].

[19] S. Gieseke, P. Stephens, and B. Webber, "New formalism for QCD parton showers," *JHEP* **0312** (2003) 045, arXiv:hep-ph/0310083 [hep-ph].

[20] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, *et al.*, "Event generation with SHERPA 1.1," *JHEP* **0902** (2009) 007, arXiv:0811.4622 [hep-ph].

[21] S. Plätzer and S. Gieseke, "Dipole Showers and Automated NLO Matching in Herwig++," *Eur.Phys.J.* **C72** (2012) 2187, arXiv:1109.6256 [hep-ph].

[22] W. T. Giele, D. A. Kosower, and P. Z. Skands, "A Simple shower and matching algorithm," *Phys.Rev.* **D78** (2008) 014026, arXiv:0707.3652 [hep-ph].

[23] T. Sjöstrand, S. Mrenna, and P. Z. Skands, "A Brief Introduction to PYTHIA 8.1," *Comput.Phys.Commun.* **178** (2008) 852–867, arXiv:0710.3820 [hep-ph].

[24] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, "Parton distributions for the LHC," *Eur. Phys. J.* **C63** (2009) 189–285, arXiv:0901.0002 [hep-ph].

[25] R. M. Duran Delgado, J. R. Forshaw, S. Marzani, and M. H. Seymour, "The dijet cross section with a jet veto," *JHEP* **1108** (2011) 157, arXiv:1107.2084 [hep-ph].

[26] G. P. Salam and G. Soyez, "A practical Seedless Infrared-Safe Cone jet algorithm," *JHEP* **05** (2007) 086, arXiv:0704.0292 [hep-ph].

[27] M. Cacciari and G. P. Salam, "Dispelling the N^3 myth for the k_t jet-finder," *Phys. Lett.* **B641** (2006) 57–61, arXiv:hep-ph/0512210. http://fastjet.fr/.

3 Tuning for Wide-Angle Radiation

3.1 Introduction

As we reach unprecedented levels of precision in experiments the uncertainties in numerical simulations need to be considered.

The results from simulations of collider physics will in general have both statistical and systematic errors, much like the experimental data. Statistical errors occur when an insufficient number of events is generated. These errors can be reduced by generating a larger set of events.

What we call systematic errors here are those which are due to the parameters of the physical models used in the event generator. Monte-Carlo event generators are reliant on physically inspired models, rather than directly on physical theory. The full physical theories may either be not completely understood, as is the case in the non-perturbative stage where partons become the final hadrons, or there may not exist an algorithm by which the full physics can be reasonably implemented. In order to provide a better representation of the experimental data these models often have variable parameters. By tuning these parameters we should be able to get a reasonable amount of predictive power.

The systematic errors can be estimated by varying the model parameters and observing how these changes manifest in relevant experimental observables. There are also other systematic errors which cannot be estimated, such as physical effects which are simply missing from the models. Additionally, uncertainties arising from the choice of parton distribution functions (PDFs) can also appear in these errors.

In Chapter 2 we discussed a modification to the parton shower which may require a tuning to be performed. We split the whole tuning process into the following steps.

- Identify which analyses are dependent on the changes we have made.

- Identify which tuning parameters will give noticeable changes to those analyses (but do not neglect their impact on others) and find an appropriate range to vary them in.
- Identify the number of events required to reasonably remove statistical errors in these analyses, taking in to account the variance of parameters.
- Generate a large number of runs with randomly chosen values for said parameters in an appropriate range.
- Use these runs to find a best fit to experimental data and take that as the new tune.
- Improve the tune by comparing to further analyses, if possible.

It is possible to tune by hand when we only have one or two parameters, but when we consider that we may have three or more parameters which are correlated this quickly becomes unfeasible. To deal with this problem, we shall use the program Professor [1].

3.2 General Tuning Considerations

3.2.1 Models within the event generator

A general event generator can be separated into multiple different parts. The first part is the hard matrix element generator, which generally runs on physically calculated quantities rather than a model. There are some ambiguities however as there are different ways of treating the quark masses.

Once the hard process has been selected by the matrix element generator it is showered by a parton shower using the DGLAP splitting kernels [2–4] with at most a few parameters. For our purposes the most important parameter in the shower is the strong coupling constant α_s. In

addition to the shower parameters it is also possible to choose a different evolution scale. If the evolution scale is correctly implemented[3] then there should be, within the accuracy of the shower, no consequence to changing between virtuality, transverse momentum or any other choice of scale.

Once the partonic phase of the evolution is complete the partons are transformed in to hadrons using a hadronization mechanism. The actual point at which the partonic evolution stops is usually a parameter of the chosen event generator. Since the behaviour contained within the hadronization mechanism is non-perturbative we must use a physically inspired model. This model will in general contain many variable parameters. These parameters will determine the multiplicities of the different variants of hadrons. Since these parameters are not fixed by theoretical methods we can vary them to obtain a better match to experimental observables. However, we expect that there will be some correlation between the parameter determining the end of the partonic evolution and some of the parameters of the hadronization mechanism.

In addition to the main hard process, there may also be additional interactions which need to be generated. These are handled by a multiple interaction model.

3.2.2 Selection of tuning parameters

If we had infinite computational time then the best approach would be to vary all possible parameters to find the best tune. However since this is not the case we need to only choose the parameters which we believe will have the largest effect on the relevant observables. Professor works by modelling the response of an event generator to changes in parameter on a bin-by-bin basis. It does so by creating a quadratic or cubic model

[3]The evolution scale is correctly implemented if it takes in to account colour coherence. If virtuality is chosen as the evolution scale then a veto on radiation outside angular ordered regions must be used [5].

of bin response to the parameter changes. The minimum number of runs, R needed to perform a tuning with cubic interpolation is [1]

$$R = 1 + N + \tfrac{1}{2}N(N+1) + \tfrac{1}{3!}N(N+1)(N+2), \tag{3.1}$$

where N is the number of parameters to vary. Given that each run may consist of over 10 million events at a rate of at most 100 events per second on a modern CPU we can see that increasing the number of parameters greatly increases the computational time required. However, to properly estimate the systematic errors we need to vary as many parameters as possible. An adequate compromise in this situation is to have 3-5 parameters. We must therefore identify the most important parameters to vary in our tune.

3.2.3 Statistics

To correctly observe the effects of changing parameters we need to minimise any error arising due to a lack of statistics. In the event generator this corresponds to generating enough events. One measure of whether we have enough events is observing how predictions change under varying the of the number of events. A simple way of doing this is to keep doubling the number of events generated. If there is no change in an observable under this doubling then the current statistics is enough for said observable. Of course the statistics required for each observable and even individual bins of the analyses may differ. In this case we need to determine which bins can be properly predicted with a reasonable number of events.

There are two approaches to generating the required statistics. The first is to simply operate with no cuts in the kinematical variables of the hard process and have the generator create events purely according to the natural weights. While this will generally give the most accurate representation of the output of the event generator it could take a very

long time to generate the required statistics. In the higher energy regions, or equivalently high rapidity and transverse momentum regions, the PDFs will be significantly smaller than in the lower energy region. In order to populate the high energy region an unreasonable amount of events will likely need to be generated.

The alternative second approach is to generate events in different kinematic slices. These kinematic slices can be of varying energies, transverse momenta or rapidities. These can then be combined either according to their individual cross sections or used in separate plots with similar kinematic requirements. This approach allows much more efficient generation of the rarer parts of phase space.

However, one needs to be careful when comparing the second approach to experimental data. This is because the cuts can only be applied to the hard process. When the shower, hadronization and multiple interaction mechanisms have been run the kinematics of the final-state particles may no longer correspond to the same cuts. Similarly, events which did not make the cut before those mechanisms may lie in the region of interest after them. Thus one should not just use cuts exactly the same as experiment, but instead with some leeway. For a thorough investigation of statistics one should also look into how varying the amount of leeway impacts on the event generator output. We should note that the upper limit of the kinematical region is usually irrelevant, as long as it contains the full region of interest. This is because any regions with larger transverse momentum or larger rapidity final state particles than the region of interest are naturally suppressed and will only form a small fraction of the events generated, assuming that the probability of such events falls with total energy.

3.2.4 Tuning procedure

The ideal tuning procedure is given in the following list:

- Generate n sets of events, each with a different set of parameters.
- Choose a set of analyses to tune to.
- Determine the event generator response to the parameter changes.
- Find the best possible tune, given the above data.

3.3 Tuning for wide-angle radiation

3.3.1 Summary of changes

The changes we have made are given in detail in both Chapter 2 and in Ref. [6]. They have been made to the parton shower in the event generator Herwig++ [7–9].

There are two changes that we have made to the colour evolution. The first is a change in how the initial scale of a gluon is chosen at the start of the parton shower. In the leading colour limit a gluon has two colour lines. Each of these lines is connected to a different parton. Hence there are two possible choices of scale for each gluon.

In the new approach the largest scale is always chosen. This ensures that if there are colour lines crossing the gap there is always a possibility of wide-angle radiation. However, it is possible that only one of the colour lines will cross the gap. In this situation, the colour factor used to generate radiation is halved to $\frac{1}{2}N_c$. Once the scale has been evolved down below that of the second colour line the original colour factor of N_c is recovered as both lines are now able to radiate. As our previous studies have shown, this results in fewer events with minimal wide-angle radiation than the previous algorithm.

The second change is that wide-angle emissions are now only connected to wide-angle colour lines. In the original algorithm each gluon emitted was connected randomly to one of the two colour lines. This resulted in incorrect behaviour when the hadronization mechanism was run. It was possible that small-angle colour lines could be connected to wide-angle radiation, resulting in a larger number of hadrons being produced in the wide-angle region. The new algorithm only assigns radiation to the lines with scales above the current evolution scale. This results in a reduction in hadrons produced in the wide-angle regions as colour lines are no longer incorrectly pulled across the gap.

While the two changes are opposite in effect we find that the first change tends to have a much greater effect than the second. From an experimental perspective it is impossible to separate out the two effects since the observed final state is always hadrons. In the generator itself we can turn off the hadronization mechanism but we shouldn't expect that the results would exactly match experimental measurements in that case.

3.3.2 Selection of relevant tuning parameters

We expect that the changes outlined above will only be noticeable in very high energy scatterings. This is because the strength of the effect is proportional to $Y\text{Log}[\frac{Q}{Q_0}]$, where Y is the size of the gap between the two hardest jets, Q is the average transverse momentum of the two hardest jets and Q_0 is the limit on the transverse momentum of any radiation between the two hardest jets. In order to minimize the impact of the underlying event a large value of Q_0 is desired. Thus, in order to maximize the effects of our changes we must have high energy scatterings in order to maximize the values for Y and Q.

With this expectation we choose to ignore the parameters of the multiple interaction model, since these additional scatterings are often at

much lower energy than the original. Thus we only need to look at the parameters of the shower and the hadronization model.

In the parton shower the main parameter that we are concerned with is the strong coupling constant α_s. We know that almost all processes will strongly depend on the coupling constant. Since the modification we have made is for QCD $2 \to 2$ scattering processes this is a parameter that should be varied.

In addition to the strong coupling constant, we should also consider varying parameters proportional to the minimum scales allowed in a splitting. These parameters define the transition between the perturbative parton shower and the non-perturbative hadronization mechanism. If we are varying parameters in the hadronization mechanism then it seems like there may be a correlation between said parameters and the minimum scales allowed. Such a correlation should be studied in order to provide the best tune possible.

In the hadronization mechanism there are a total of 25 parameters and 5 switches. Since we are mainly interested in QCD $2 \to 2$ scattering processes we do not really need to consider the various flavours of the hadrons and quarks. As a result the above can be reduced to 7 parameters and 3 switches.

The fastest way to gauge the effects of varying the parameters is to choose an observable which is highly dependent on the wide-angle radiation. To that end we chose the gap fractions for average transverse momenta of the bounding jets greater than 200 GeV and rapidity gaps in the region $4 \leq Y \leq 5$. The gap fraction is the fraction of events which do not contain any jets with transverse momentum greater than Q_0 between the two bounding jets. While higher transverse momenta and rapidity gaps can be reached at the LHC there is not sufficient statistics to accurately see how varying parameters changes the results.

The parameters were typically varied at their extremes and at some points in the middle of the possible ranges.

After the tests we found that the five most important parameters are `AlphaMZ`, `cutoffKinScale`, `PSplitLight`, `ClPowLight`, and `ClMaxLight`.

3.3.3 Interpretation of parameters

The parameter `AlphaMZ` corresponds physically to the strong coupling constant at the scale of the mass of the Z boson. This controls the overall normalization of the coupling constant. Increasing the coupling constant results in more radiation and hence more jets. Since this affects almost all analyses we should only let it vary a small amount.

The parameter `cutoffKinScale` is related to the minimum allowed virtuality of the gluon. The hadronization mechanism in Herwig++ requires that all gluons at the end of the shower are decayed in to quark and anti-quark pairs. Since quarks are treated as having masses this requires that the gluon have a minimum virtuality at the end of the shower in order for it to be able to decay. The minimum gluon virtuality Q_g is defined as

$$Q_g = \max\left(\frac{\delta - am_q}{b}, c\right), \tag{3.2}$$

where δ is `cutoffKinScale`, m_q is the mass of a quark and a, b and c are tuning parameters. For light mass quarks we can therefore see that Q_g is proportional to `cutoffKinScale`.

The factors `ClPowLight`, `ClMaxLight` and `PSplitLight` are used in the hadronization mechanism. When the perturbative shower finishes what remains is a set of coloured partons. All of the gluons are decayed into quark and anti-quark pairs. The resultant partons then form colour singlet clusters by pairing colour connected quark and anti-quark pairs. However, these clusters can have masses orders of magnitudes higher than most common hadrons. The mass of a cluster M formed by partons with

momenta $p_{1,2}$ and masses $m_{1,2}$ is

$$M = (p_1 + p_2)^2. \tag{3.3}$$

These clusters are decayed in the model if

$$M^{Cl_{pow}} \geq Cl_{max}^{Cl_{pow}} + (m_1 + m_2)^{Cl_{pow}}. \tag{3.4}$$

From this expression we can see that if we increase `ClPowLight` we will get more clusters produced as M always has the greatest value. Similarly, a smaller `ClMaxLight` can increase the number of clusters produced, which results in a greater spread of resulting hadrons and possibly more jets.

When a cluster decays it will split into two more clusters. This is done by either popping a quark and antiq-uark pair out of the vacuum or occasionally a diquark and anti-diquark pair. Assuming that there is the quark and anti-quark pair popped from the vacuum then the masses of the new clusters clusters $M_{1,2}$ produced in the splittings are given by

$$M_1 = m_1 + (M - m_1 - m_q)R_1^{1/P}, \tag{3.5}$$

$$M_2 = m_2 + (M - m_2 - m_q)R_2^{1/P}, \tag{3.6}$$

where m_q is the mass of the quark popped from the vacuum to form the new cluster. Here P refers to `PSplitLight` and the R_i are random numbers in the range 0 to 1. Decreasing `PSplitLight` results in higher mass clusters being produced and therefore more splittings. This results in more hadrons, but each of the hadrons will have lower energy.

Once the clusters reach low enough masses they become the hadrons which are observed in the detector and the hadronization mechanism terminates.

Parameter	Default	Min	Max
`AlphaMZ`	0.12	0.11	0.13
`cutoffKinScale`	2.65	2.00	5.00
`PSplitLight`	1.20	0.50	3.50
`ClPowLight`	1.28	0.50	3.50
`ClMaxLight`	3.15	0.50	3.50

Table 1: Default values for parameters and the range of variance for tunes.

3.3.4 Parameter ranges

We should look at the possible ranges for the values of said parameters and their effects upon observables. To look at the observable response we vary the parameters at their extrema and around the default. The default parameter values in the current Herwig++ tune and the regions used in the tune are given in Table 1. The results for the five parameters are shown in Figures 24-28.

Figure 24 shows the variance of generator output when the parameter `AlphaMZ` is varied between 0.100 and 0.135. The end-points are chosen by the limits of the event generator. At low Q_0 there is quite a large dependance on the value of the strong coupling.

Figure 25 shows the variance of generator output when the parameter `cutoffKinScale` is varied between 0.5 GeV and 10.0 GeV. The distribution is fairly independent of the gluon virtuality cut up to a value of 5.0 GeV. At a value of 10.0 GeV the gap fraction is universally higher than the other values. We would expect that there should be no difference in results from this range of variance as `cutoffKinScale` is always lower than Q_0. The different results observed at a value of 10.0 GeV is likely due to less emissions at the edges of the gap, which results in fewer jets in the gap and therefore a higher gap fraction. The experimental results are matched fairly well within the range bound by the other values.

In Figure 26 we show the generator response to varying the parameter `PSplitLight` from a value of 0.1 to a value of 9.9. For small values the

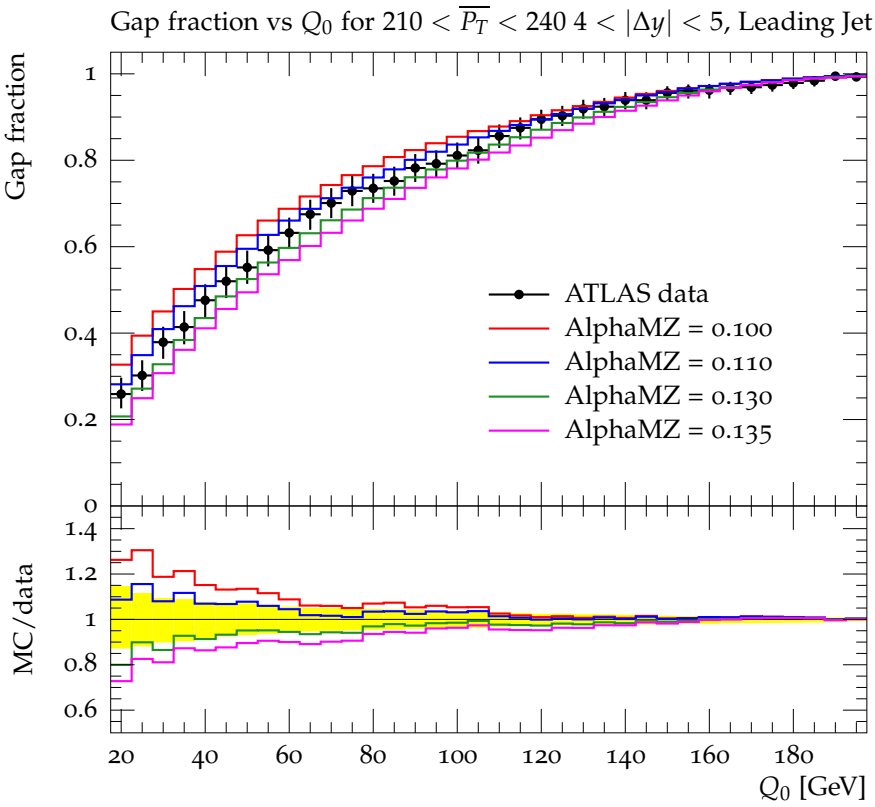

Figure 24: Gap fraction as a function of veto scale Q_0 for the kinematical region $210 \leq Q \leq 240$ and $4 \leq |\Delta y| \leq 5$. The rapidity gap is bounded by the two hardest jets. The points are ATLAS data [10]. The histograms show the generator output for varying values of `AlphaMZ`.

distribution is roughly the same. However, as we increase the value to the upper limit set by the generator we find that we are now unable to match the experimental results. Hence we must constrain the values used in the tune to smaller values.

Next we show Figure 27 which contains the generator response to varying the parameter `ClPowLight` from a value of 0.1 to 9.9. When varying the parameter over large values the distribution is unchanged, except at very low values. Hence we need to set a limit on the lowest possible value allowed in the tune.

Finally we show Figure 28 which contains the generator response to varying the parameter `ClMaxLight` from a value of 0.1 to 9.9. The

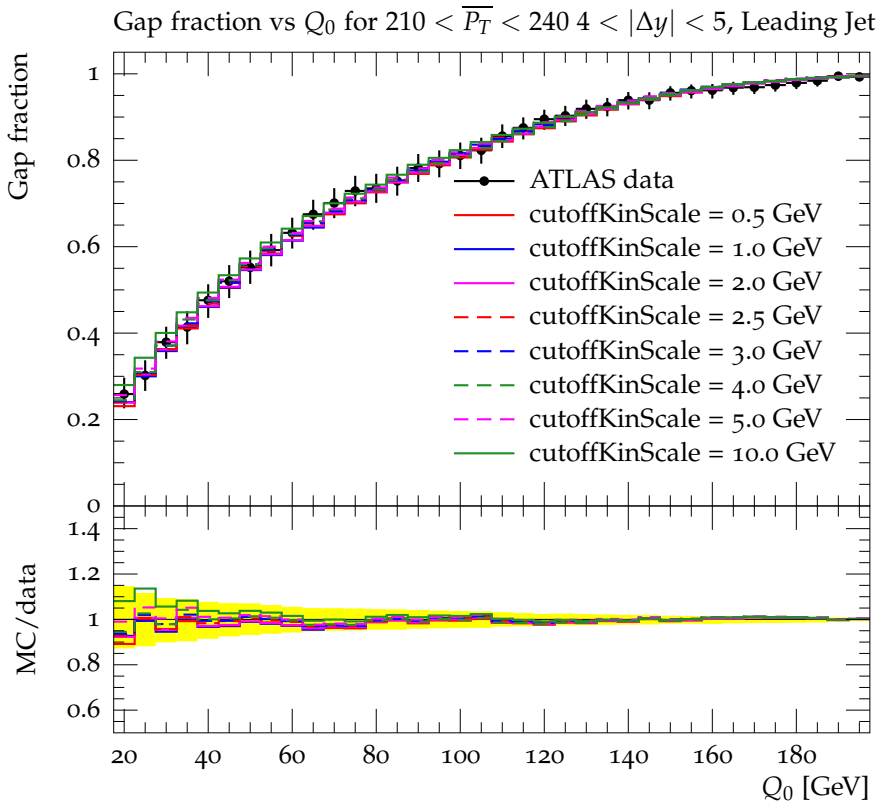

Figure 25: Gap fraction as a function of veto scale Q_0 for the kinematical region $210 \leq Q \leq 240$ and $4 \leq |\Delta y| \leq 5$. The rapidity gap is bounded by the two hardest jets. The points are ATLAS data. The histograms show the generator output for varying values of `cutoffKinScale`.

distribution is unchanged under the varying of this parameter. However, we will include this parameter in the tune as it appears in the same equation as `ClPowLight`.

While this approach allows us to observe the effects of varying single parameters, it does not suggest how we could change all five parameters in order to get the best tune. We will need a more systematic approach in order to do this. This is what is provided by the program Professor.

3.3.5 Selection of relevant analyses

Only observables sensitive to wide-angle radiation will be affected by our modification. Additionally, since our changes have only been made to the

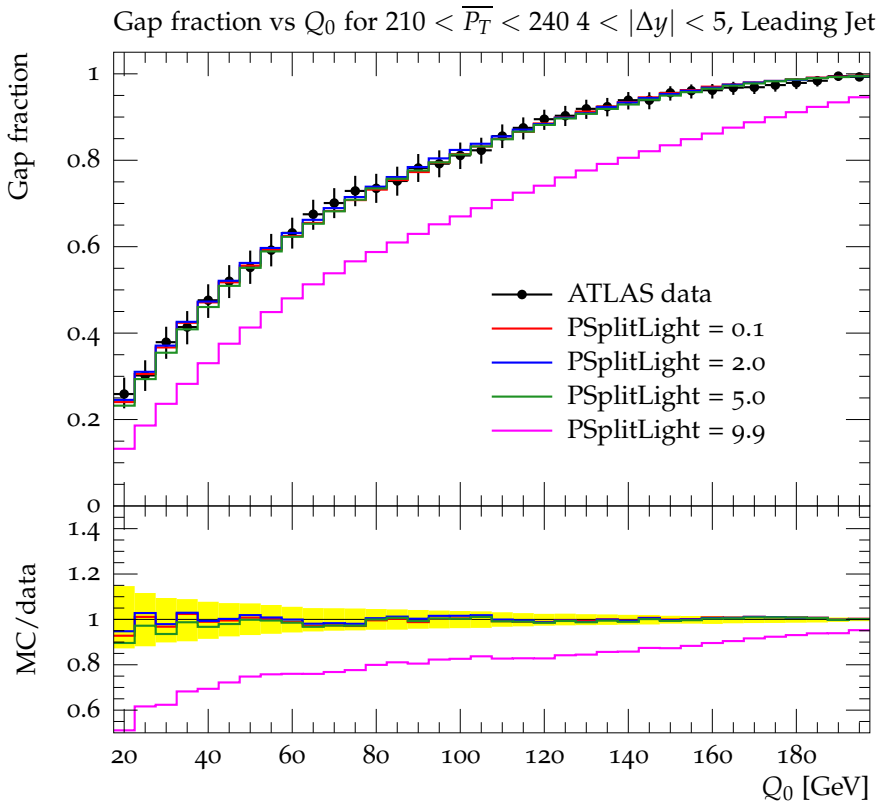

Figure 26: Gap fraction as a function of veto scale Q_0 for the kinematical region $210 \leq Q \leq 240$ and $4 \leq |\Delta y| \leq 5$. The rapidity gap is bounded by the two hardest jets. The points are ATLAS data. The histograms show the generator output for varying values of `PSplitLight`.

radiation from hard gluons, we can rule out any observable which cannot be generated with gluons in the hard process. Thus we should mainly consider QCD $2 \to 2$ scattering processes at the LHC. From previous simulations we have found that the energies at the Tevatron are not sufficient to see a difference between the two approaches. Hence we will concentrate mainly on the currently available LHC data.

The main analyses that we are considering are the measurements of dijet production with a veto on additional central jet activity [10] and dijet azimuthal decorrelations [11, 12]. In the ATLAS measurement of dijet production with a veto, there is a restriction applied to the amount of radiation allowed in the region between either the two hardest jets or

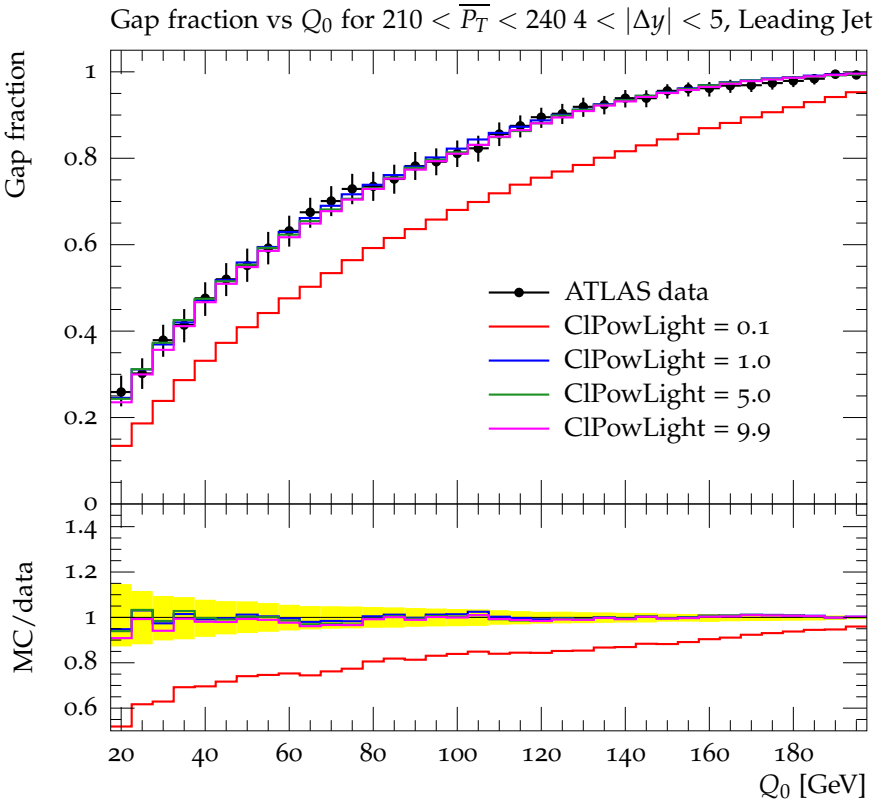

Figure 27: Gap fraction as a function of veto scale Q_0 for the kinematical region $210 \leq Q \leq 240$ and $4 \leq |\Delta y| \leq 5$. The rapidity gap is bounded by the two hardest jets. The points are ATLAS data. The histograms show the generator output for varying values of `ClPowLight`.

the most forward and most backward jet. We define the hardest jets as those with the highest transverse momentum and the most forward and backward jets as those with the highest and lowest rapidity respectively[4].

At the lowest order in QCD there will only be two jets in the process. Thus this observable is directly dependent on the implementation of QCD radiation in the parton shower. Wide-angle radiation from the hard process has a chance of landing in the gap between the two jets. It is also possible for small-angle radiation to just make it into the gap.

In both Chapter 2 and in Ref. [6] we have shown that the gap fraction

[4]Since rapidity is not positive definite when we say lowest rapidity we actually mean most negative.

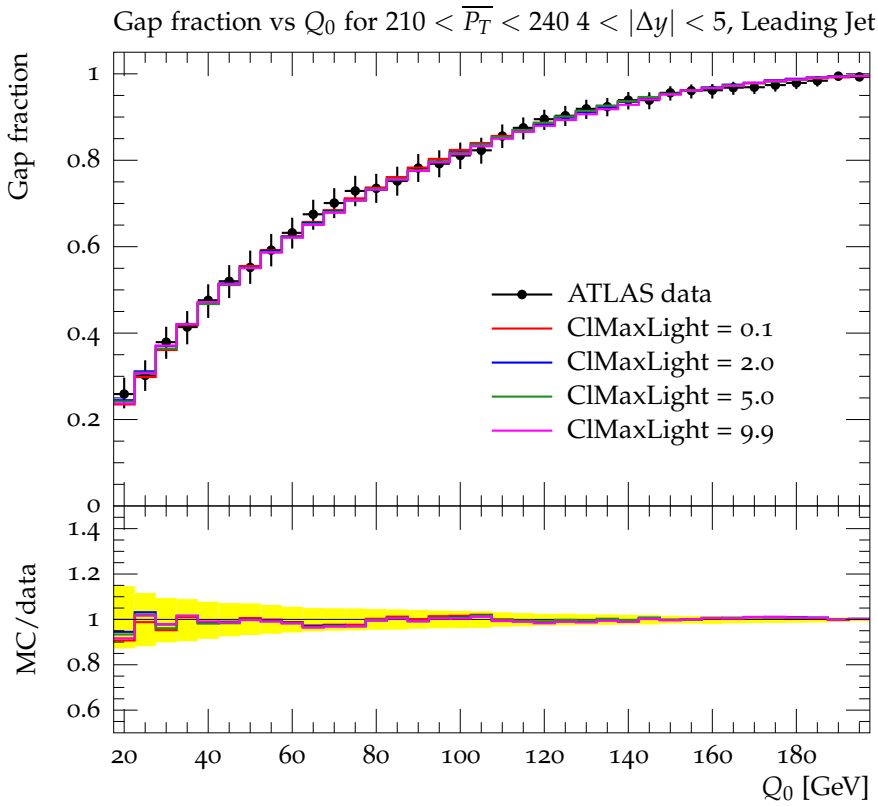

Figure 28: Gap fraction as a function of veto scale Q_0 for the kinematical region $210 \leq Q \leq 240$ and $4 \leq |\Delta y| \leq 5$. The rapidity gap is bounded by the two hardest jets. The points are ATLAS data. The histograms show the generator output for varying values of `ClMaxLight`.

as a function of the veto scale is very sensitive to the form of the evolution, at least for pure gluon processes at 14 TeV. This observable was described at length in Chapter 2.

Next we consider the other main observable which is dijet azimuthal decorrelations. These analyses at both ATLAS and CMS look at the difference in angle between the two hardest jets in an event. At the lowest order in QCD the two jets should be back-to-back. Thus if we see any departure from this we find that it is entirely due to radiation from the hard process. Any wide-angle radiation will cause a difference between the lowest order result and that which is experimentally measured. Importantly we should note that while the parton shower generates much

of the softer radiation correctly, it will not generate the correct behaviour when it comes to hard jets. Thus even a re-tune of the parton shower and hadronization mechanism should not be able to neccesarily match the experimental results for very hard three-jet configurations.

In addition to the main analyses, we can also include an additional set of analyses detailed in Refs. [13–18] since these analyses can run at the same time as those above for minimal computational cost.

The comparison of Herwig++ results to the experimental data is handled by the program RIVET [19].

3.3.6 Statistics

To perform the tuning we will split our runs into sets of different energy regions. This will allow us to generate the required statistics for the analyses of choice with a reasonable computational time. In terms of the actual data sets we have decided that running one million events in each of the energy regions offers the best convergence to time ratio and also allows us to merge different energy sets in a much easier way.

Figure 29 shows the convergence of statistics for a gap fraction in the high energy region. Here we choose a kinematic slice with $Q \geq 200\text{GeV}$ and no restriction on rapidity. This region is hard to populate as it has both a high average transverse momentum and a large rapiditity gap. For event numbers of less than one million the event generator prediction is unstable and dependent on the number of runs and the seed of the random number. Above one million events run there is no change in behaviour. Hence running more than one million events is unnecessary for this observable.

For this observable we could have obtained the same result with fewer events if we had restricted the rapidity region to approximately that of the experimental analysis. This would be the most convenient method if

we only had one analysis to consider. However, when we have multiple analyses we would have to perform a run for each of the analyses with the correct kinematical slice. In addition to this, there may be analyses that do not choose a specific rapidity. For these we would have to find the correct way to combine the results of the different slices.

Due to these added complications we choose to only restrict one kinematic variable in the hard process as it makes merging different slices much simpler. Which method ultimately requires less statistics to generate the correct output will need further study.

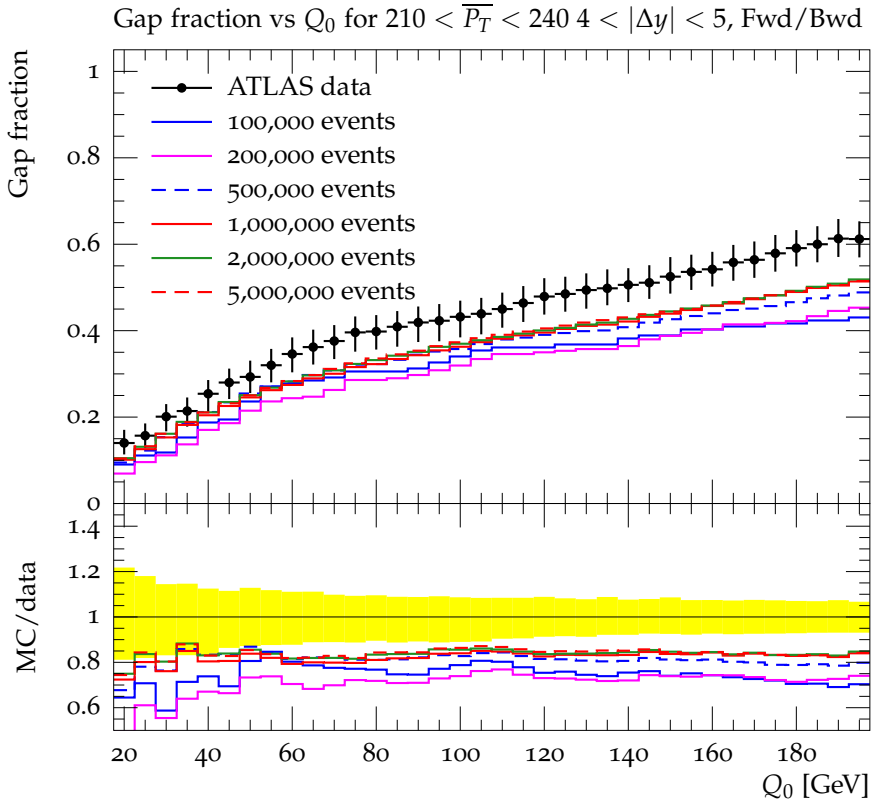

Figure 29: Gap fraction as a function of veto scale Q_0 for the kinematical region $210 \leq Q \leq 240$ and $4 \leq |\Delta y| \leq 5$. The rapidity gap is bounded by the most forward and backward jets. The points are ATLAS data. The histograms are Herwig++ responses with different statistics ranging from 100,000 events to 5,000,000 events.

3.3.7 Tuning approach

Having decided upon the variables required and analyses which are most important we now need to begin our tuning. The first step is to generate all of the data needed for the tuning. Since we are looking at potentially billions of events, the only logical place that can provide the required computational power is the LHC computing grid. Our final data set is 80 choices of parameter with 23 million events each.

3.3.8 Improving with additional data

While we are mainly focused on LHC observables, we should also look at how the Tevatron and e^+e^- results vary with our tunes. While most of the Tevatron and e^+e^- results were unchanged by our modification to the colour evolution of the parton shower they may not remain unchanged under a tuning. We can also go a step further and attempt to tune to observables at these two colliders.

3.4 Tunes

We currently have the event generator's responses for 80 sets of parameters for both LHC and e^+e^- observables. The observables used in each run are given in Table 2. For all of these runs we have turned on both hadronization and multiple interactions. In order to have a cubic interpolation for 5 different parameters we need the event generator's responses for at least 56 sets of parameters. As we have 80 event generator responses we should be able to use this cubic interpolation to provide the best fit.

The PDF set used is the default Herwig++ choice of MRST [20]. Since we only used one PDF set for our tuning we are unable to estimate the full systematic errors due to the choice of PDF.

Analysis	Observable	Tune						
		1	2	3	4	5	6	7
Ref. [10]	All	1.0	1.0	1.0	1.0	1.0	0.0	1.0
Ref. [10]	(d04-5)-*	1.0	0.0	0.0	0.0	0.0	0.0	0.0
Ref. [10]	*-y02	1.0	1.0	0.0	1.0	1.0	0.0	1.0
Ref. [10]	(d19-25)-*	1.0	1.0	0.0	1.0	1.0	0.0	1.0
Ref. [10]	(d30-36)-*	1.0	1.0	0.0	1.0	1.0	0.0	1.0
Refs. [11, 12]	All	0.0	0.0	0.0	1.0	1.0	0.0	1.0
Ref. [11]	*-(y08-09)	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Refs. [13–18]	All	0.0	0.0	0.0	0.0	1.0	0.0	1.0
Refs. [21–33]	All	0.0	0.0	0.0	0.0	0.0	1.0	1.0

Table 2: Observable weights for the various tunes. The symbol * denotes the sum of all possible observables matching that pattern. Brackets indicate the sum of the range of analyses with.

3.4.1 Tune 1: Gaps Between Jets with no bins excluded

In this tune we focus on only the main observable, which is the gaps between jets measurements. This tuning should provide the best possible match to the observable that is of greatest interest to us. Since there are quite a few analyses with varying transverse momentum and rapidity regions this should also give a fairly good general tune for wide-angle radiation.

3.4.2 Tune 2: Gaps Between Jets with higher bins excluded

Since we had to limit our statistics in the generated runs we do not believe that we will be able to match some of the experimental data. Additionally, we believe that it may not be possible to describe some of the higher energy regions as these are where corrections due to hard jets become important. Furthermore, some of the experimental data seems to be extremely limited by statistics in the high energy regions. As such we choose to tune only to the bins that we believe we have a reasonable chance of predicting.

3.4.3 Tune 3: Gaps Between Jets without forward and backward observables

It is not only statistics that limit our ability to predict observables. The lack of hard radiation generated by the parton shower makes it unreliable for predicting observables where the bounding jets are not those of the hard process. In this tune we choose to match to analyses where the bounding jets are those which correspond to the highest transverse momentum jets only.

3.4.4 Tune 4: Gaps Between Jets and Dijets

While tuning to only one experimental analysis will give the best fit to that set of observables, we must be careful that we do not obtain a tune which does not fit any other data. As such we must consider simultaneously tuning to other observables. The other two most important experimental analyses were the dijet azimuthal decorrelation measurements. As noted in the previous section, these analyses also contain regions that we are unable to correctly predict, either due to a lack of statistics or missing high energy jet physics.

3.4.5 Tune 5: All LHC analyses with some bins excluded

Rather than just matching to selective analyses, we should check to see what happens when we tune to all possible relevant observables. In this tune we attempted to match to all the possible LHC analyses for QCD $2 \to 2$ scattering processes.

3.4.6 Tune 6: Pure e^+e^- Tune

In this tune we attempt to match our event generator output to only e^+e^- annihilation observables. It is believed that this should give a similar tune to the current one, since we have not modified any behaviour which occurs

in e^+e^- events. Of course, there are many different ways to tune. The final results will depend on the choice of weights for different observables.

The analyses that we are tuning to are detailed in Refs. [21–33].

3.4.7 Tune 7: e^+e^- and LHC Tune

Extending our methodology from Tune 5 we can attempt to simultaneously tune to both LHC and e^+e^- data.

3.5 Results

Having discussed the motivation for our different tunes we will now discuss the results for said tunes.

3.5.1 Correlations

While some of the parameters we are tuning should be independent, it is possible that there are correlations. What we have found is that the correlations are highly dependent on the choice of analyses that we tune to. In Figure 30 we show the correlations for Tune 4. In this case there are a number of correlations and anti-correlations between most of the parameters. From this we might conclude that tuning to specific physical processes can lead to lots of correlations between the parameters.

Next we have Figure 31 which shows the correlations for Tune 6. We see that by tuning to a different set of observables, this time e^+e^- processes, we have a different set of correlations. Some are still the same, such as the strong correlation between `ClMaxLight` and `ClPowLight` which occur in the same equation. Many correlations have become anti-correlations and vice versa.

Finally, we have Figure 32 which shows the correlations for Tune 7. In this tune we match to both LHC and e^+e^- processes. By including a larger set of observables the majority of the correlations have disappeared.

While some parameters may be correlated on one observable, they may have no correlation or be anti-correlated on another, as observed with the previous tunes. Only the parameters which are strongly correlated in all cases remain correlated for this tune.

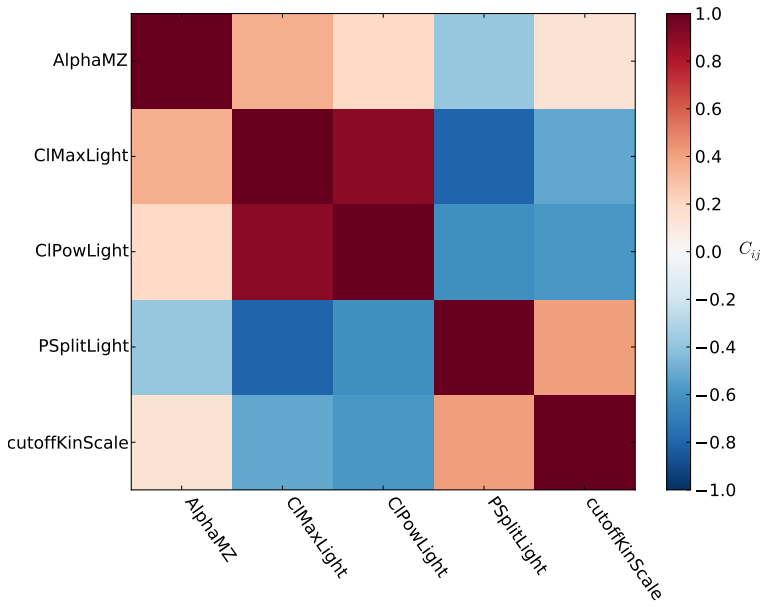

Figure 30: Correlations between parameters of Tune 4 for the minimised results. Highly correlated parameters have values close to one, uncorrelated parameters have a value of zero and anti-correlated parameters have values close to minus one.

3.5.2 Tune values

The parameter values for the tunes outlined above are given in Tables 3 and 4. The χ^2 values are shown for Refs. [10–18].

We can clearly see that each of these tunes gives a fairly different set of parameters and they all tend to be different than the current tune. The values of α_s chosen in the tunes spans a much smaller range than that allowed in the tuning process. Clearly $\alpha_s \sim 0.12$ is preferred for the current data.

The first two tunes, which differ only in the removal of a few high

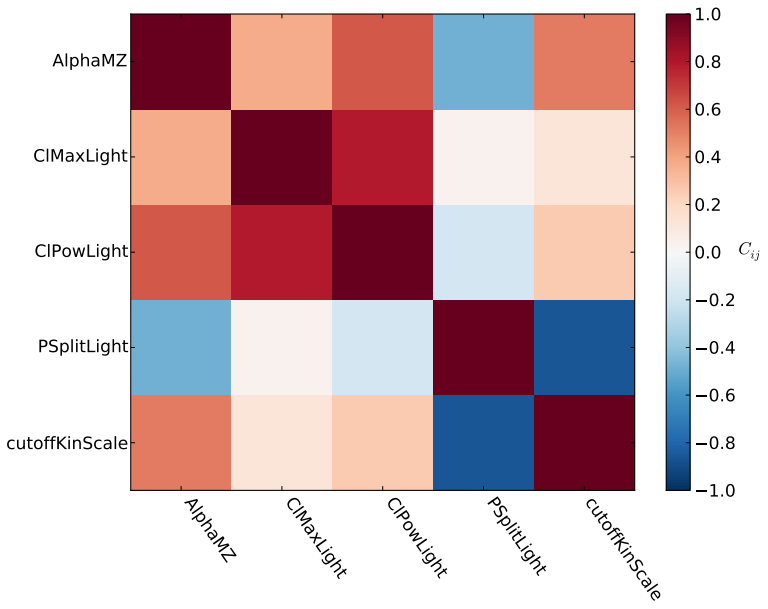

Figure 31: Correlations between parameters of Tune 6 for the minimised results. Highly correlated parameters have values close to one, uncorrelated parameters have a value of zero and anti-correlated parameters have values close to minus one.

energy observables, are extremely similar. The difference between tunes 5 and 7 is also especially notable. The inclusion of e^+e^- data has shifted the combined tune to be extremely similar to that of pure e^+e^-. Additionally, this full tune is now very similar to that of the default tune, at least in the shower parameters.

It it clear from the χ^2 values that, for a general fit, the default implementation without the colour evolution modification is the best choice. The best tunes with colour evolution for a general fit are the default tune and tune 5. We would expect that tune 5 would give the best overall result for LHC observables as these are what it is matched to. By matching to LEP observables as well we find that tune 7 gives worse results for observables than tune 5 in general.

The first three tunes only do well for predictions of observables in Refs. [10] and [15]. For tune 3 we can see that, by removing matching to

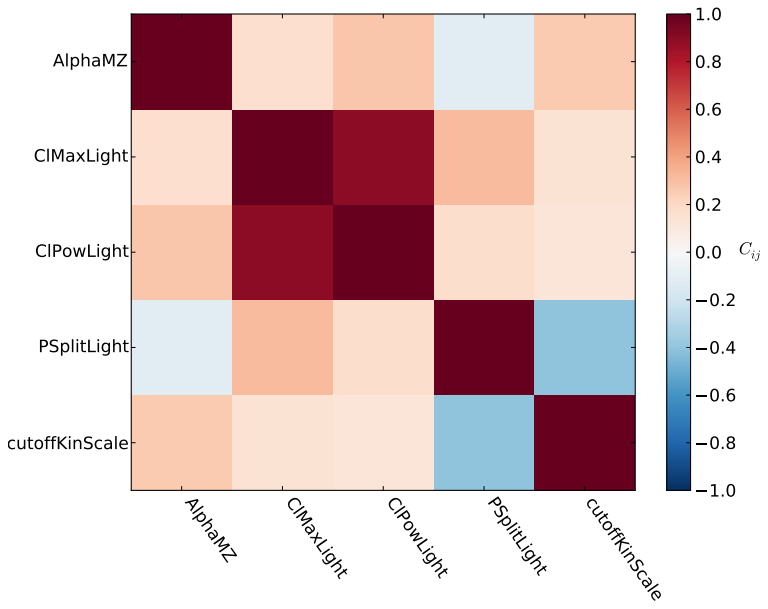

Figure 32: Correlations between parameters of Tune 7 for the minimised results. Highly correlated parameters have values close to one, uncorrelated parameters have a value of zero and anti-correlated parameters have values close to minus one.

forwards and backwards observables that cannot be predicted accurately by the parton shower, the results are generally better for most observables.

The Herwig++ event generator as a whole, regardless of tune, does particularly poorly for Refs. [13], [14] and [18]. In Ref. [13] the observables involve both large energies and configurations with three hard jets, which are problematic for the parton shower. For Ref. [14] there seems to be an issue with the overall normalization for the cross sections. The actual shape of the observables is still predicted fairly well. Finally, Ref. [18] does extremely poorly in all cases. This is again due to higher numbers of hard jets as well as high energies.

3.5.3 Resulting plots

In Figure 33 we show the results for the gap fraction in a high energy event region where the bounding jets are chosen to be the two hardest in

Parameter	Default	Tune 1	Tune 2	Tune 3	Tune 4
`AlphaMZ`	0.120	0.118	0.118	0.121	0.122
`cutoffKinScale`	2.65	4.61	4.35	4.00	3.18
`PSplitLight`	1.20	3.34	3.47	3.47	2.01
`ClPowLight`	1.28	1.39	1.56	2.51	1.89
`ClMaxLight`	3.15	1.81	1.99	2.27	2.71
χ^2 `/ d.o.f.`					
Ref. [10]	5.51	3.90	3.90	4.23	4.66
Ref. [10] (No F/B)	2.38	3.11	3.14	2.96	3.24
Ref. [11]	3.29	14.7	14.8	12.7	10.8
Ref. [12]	3.11	25.6	26.8	21.2	17.8
Ref. [13]	11.6	21.3	21.4	19.6	18.0
Ref. [14]	30.1	31.3	32.5	32.6	29.9
Ref. [15]	1.24	0.74	0.80	0.76	1.44
Ref. [16]	8.56	585	660	606	53.6
Ref. [17]	5.81	48.5	48.0	38.4	30.6
Ref. [18]	2642	2431	2474	2503	2586

Table 3: Parameters and fit quality for the various tunes. The default tune is the normal tune used for Herwig++ 2.5.2. The default fit values are without the colour evolution modification. The χ^2 values are calculated by comparing to the observables with the weights from tune 5, except for Ref. [10] (No F/B) which uses the weights from tune 3. The weights for these tunes are given in Table 2.

the event. We can see that regardless of the tune we are able to match the experimental data quite well. The default tune is the worst at predicting this observable. Tunes 1-5 and 7 are all matched to this analysis. Tune 6 is independent of this analysis.

Next, we have Figure 34 where we show the results for the gap fraction in a high energy event region where this time the jets bounding the gap are those which are the most forward and backward. Here we can see that the results are dependent on the tune of the generator. We see that it is not possible with our current tuning methodology to match the experimental results. Since almost 40% of the events in this region of phase space have a central jet with transverse momentum greater than 200 GeV it is

Parameter	Default	Tune 5	Tune 6	Tune 7	Default (E)
`AlphaMZ`	0.120	0.121	0.118	0.117	0.120
`cutoffKinScale`	2.65	3.56	2.47	2.60	2.65
`PSplitLight`	1.20	1.57	1.30	1.26	1.20
`ClPowLight`	1.28	0.730	0.683	0.621	1.28
`ClMaxLight`	3.15	3.28	2.36	2.12	3.15
χ^2 `/ d.o.f.`					
Ref. [10] (All)	5.51	4.40	4.48	4.30	4.72
Ref. [10] (No F/B)	2.38	3.11	3.29	3.44	3.16
Ref. [11]	3.29	11.8	12.3	13.2	11.0
Ref. [12]	3.11	18.8	21.0	22.8	17.6
Ref. [13]	11.6	18.7	19.4	20.2	18.3
Ref. [14]	30.1	29.5	30.7	31.5	31.5
Ref. [15]	1.24	1.14	2.48	2.79	2.45
Ref. [16]	8.56	12.8	8.03	7.76	8.51
Ref. [17]	5.81	33.2	34.5	38.2	29.6
Ref. [18]	2642	2600	2636	2628	2628

Table 4: Parameters and fit quality for the various tunes. The default tune is the normal tune used for Herwig++ 2.5.2. The default fit values are without the colour evolution modification. The default (E) fit values are with the colour evolution modification. The χ^2 values are calculated by comparing to the observables with the weights from tune 5, except for Ref. [10] (No F/B) which uses the weights from tune 3. The weights for these tunes are given in Table 2.

perhaps not too surprising that this observable is not predicted well by the parton shower. In order to correctly match this observable hard three-jet events must be generated prior to the parton shower. When Herwig++ is combined with POWHEG [34] this observable can be predicted with much better accuracy [10].

We find that the default tune gives the worst results. We can see that the default tune without the colour evolution modification is particularly bad at predicting the observables with forwards and backwards boundary jets by look at Tables 3 and 4. When these observables are removed the χ^2 value reduces from 5.51 to 2.38, which is lower than that of any of

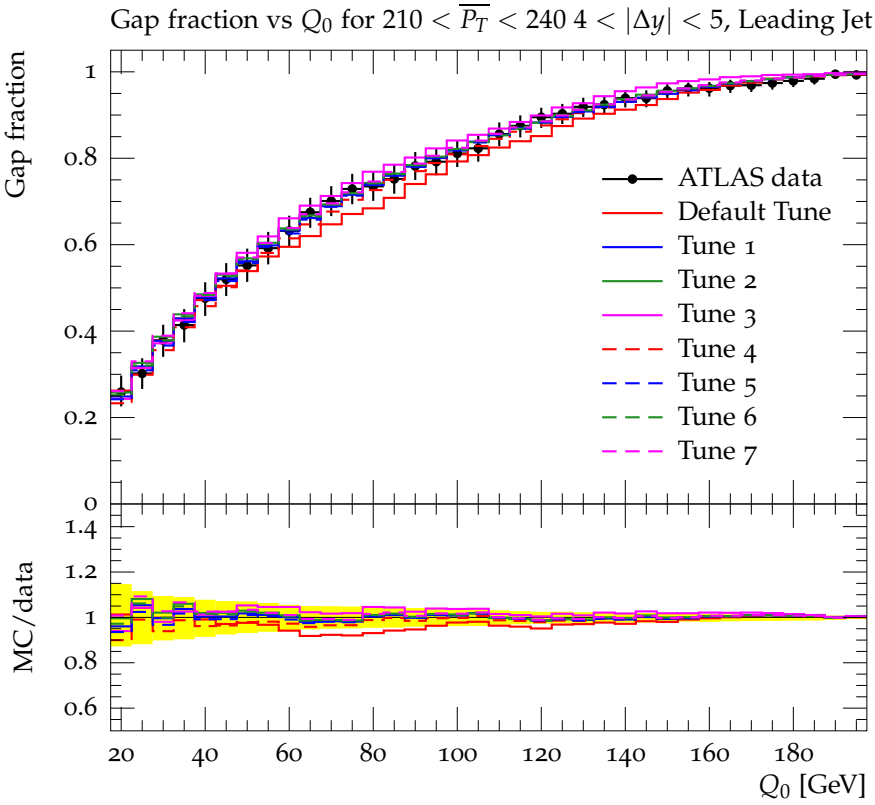

Figure 33: Gap fraction as a function of veto scale Q_0 for the kinematical region $210 \leq Q \leq 240$ GeV and $4 \leq |\Delta y| \leq 5$. The rapidity gap is bounded by the two hardest jets. The dots represent experimental data and the histograms represent the generator responses with the specified tunes. The new tunes are with the colour evolution modification while the default tune is without said modification.

the other tunes. Tunes 1-2, 4-5 and 7 are all matched to this analysis. Tunes 3 and 6 are independent of this analysis.

Our third plot, Figure 35, shows the results for the dijet azimuthal correlations in the region 200 GeV $\leq Q \leq$ 300 GeV. It is clear that we are still unable to match the results of experiment on this observable. However, we find that in this case the default tune is still the closest. Of our tunes it is difficult to tell which performs best. Tunes 4-5 and 7 are all matched to this analysis. Tunes 1-3 and 6 are independent of this analysis.

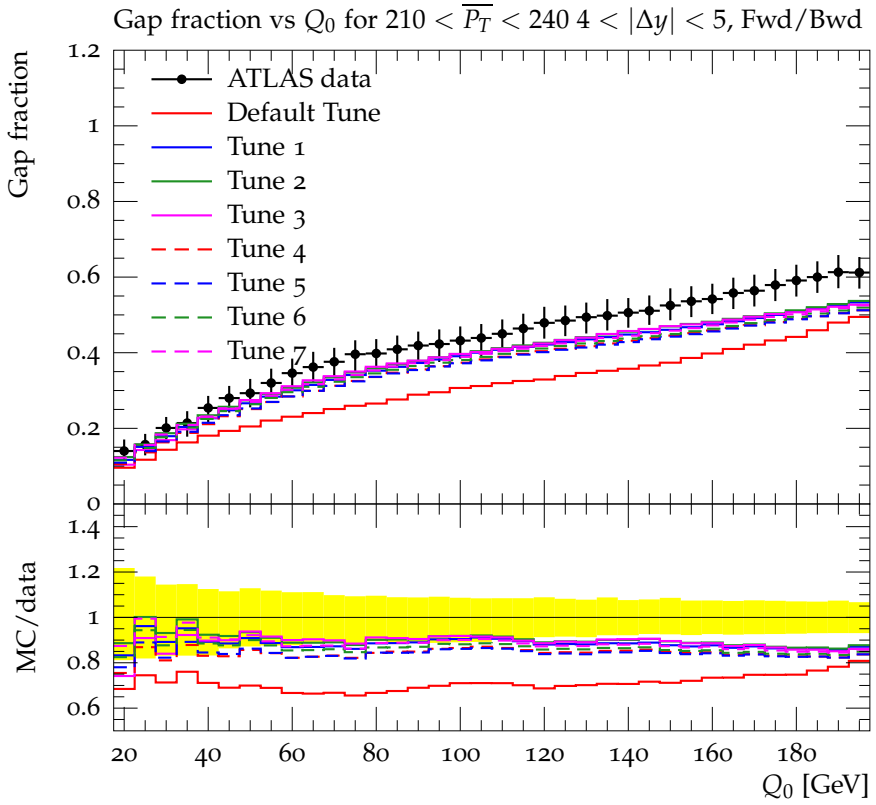

Figure 34: Gap fraction as a function of veto scale Q_0 for the kinematical region $210 \leq Q \leq 240$ GeV and $4 \leq |\Delta y| \leq 5$. The rapidity gap is bounded by the most forward and backwards jets. The dots represent experimental data and the histograms represent the generator responses with the specified tunes. The new tunes are with the colour evolution modification while the default tune is without said modification.

Our final plot, Figure 36, shows the charged hadron multiplicity as a function of rapidity. We can see that the accuracy of the event generator is highly dependent on the choice of tune. Tunes 5 and 7 are both matched to this analysis. Tunes 1-4 and 6 are independent of this analysis. Tunes 1-3 do especially poorly in this case.

3.6 Conclusion

With the ever increasing precision of both analytical calculations and experimental data it is important that we understand the errors involved

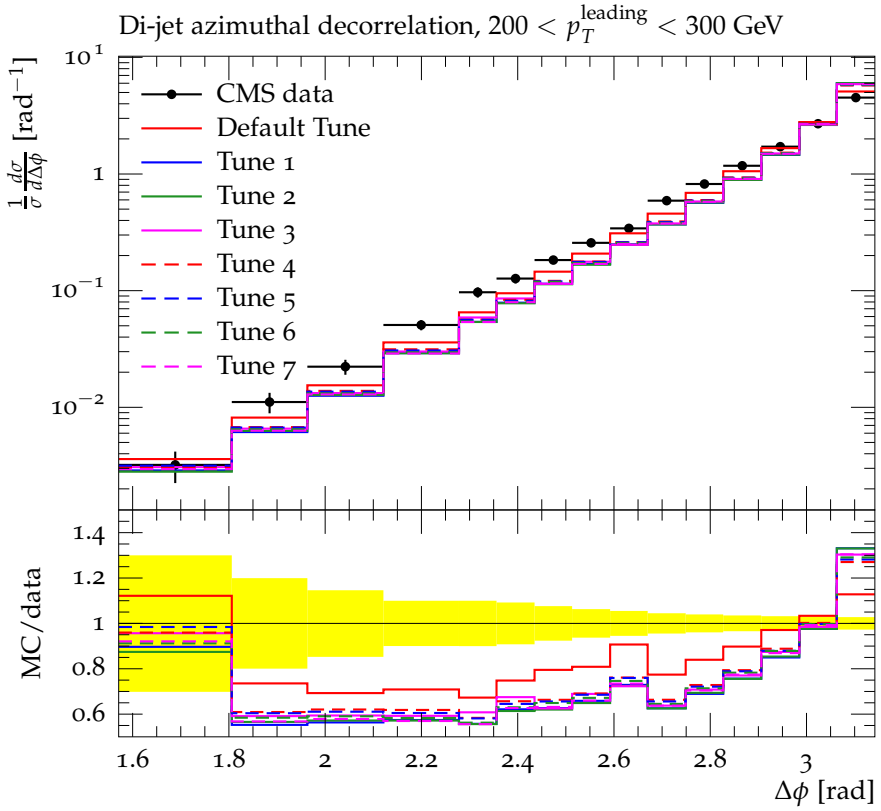

Figure 35: Dijet azimuthal correlations in the region 200 GeV $\leq Q \leq$ 300 GeV. The dots represent experimental data and the histograms represent the generator responses with the specified tunes. The new tunes are with the colour evolution modification while the default tune is without said modification.

in our event generators. By looking at the generator responses to varying model parameters we can get a good handle on the possible systematic errors for any prediction.

It is important that whenever we add new physics into one or more of the models of an event generator we perform a re-tune of the relevant parameters.

Of course, it is not possible to make complete predictions when we have incomplete models within our event generators. In the next chapter we will discuss how we can improve the parton shower by adding some of the currently missing physics.

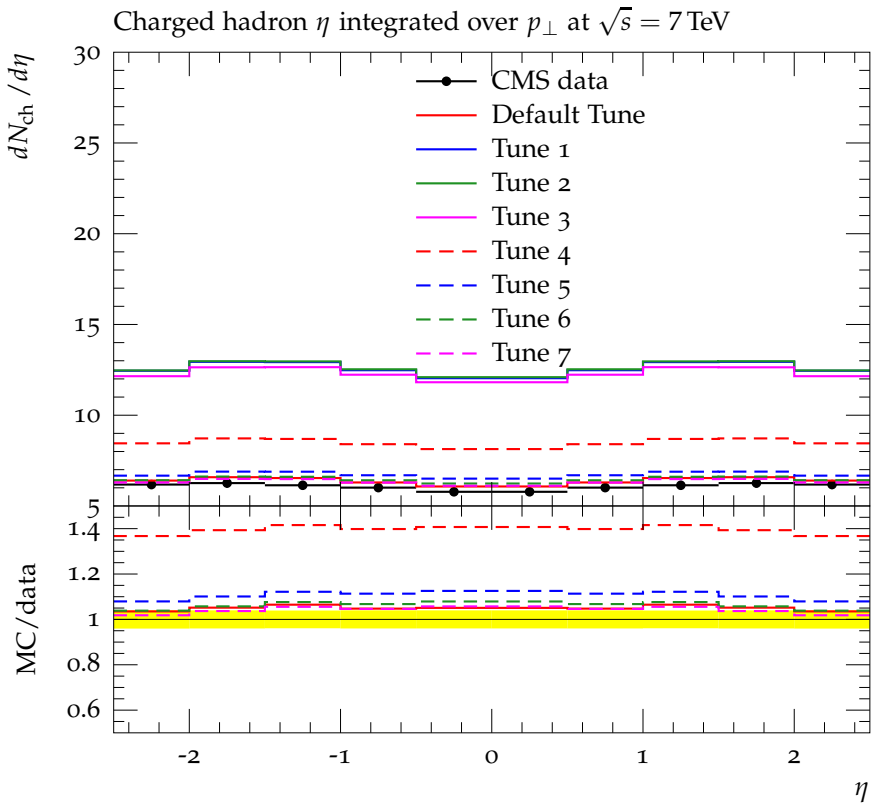

Figure 36: Charged hadron multiplicity as a function of rapidity. The dots represent experimental data and the histograms represent the generator responses with the specified tunes. The new tunes are with the colour evolution modification while the default tune is without said modification. The histograms for Tunes 1 and 2 overlap.

References

[1] A. Buckley, H. Hoeth, H. Lacker, H. Schulz, and J. E. von Seggern, "Systematic event generator tuning for the LHC," *Eur.Phys.J.* **C65** (2010) 331–357, `arXiv:0907.2973 [hep-ph]`.

[2] V. N. Gribov and L. N. Lipatov, "Deep inelastic e p scattering in perturbation theory," *Sov. J. Nucl. Phys.* **15** (1972) 438–450.

[3] G. Altarelli and G. Parisi, "Asymptotic Freedom in Parton Language," *Nucl. Phys.* **B126** (1977) 298.

[4] Y. L. Dokshitzer, "Calculation of the Structure Functions for Deep Inelastic Scattering and e^+ e^- Annihilation by Perturbation Theory in Quantum Chromodynamics," *Sov. Phys. JETP* **46** (1977) 641–653.

[5] **CDF Collaboration** Collaboration, F. Abe *et al.*, "Evidence for color coherence in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV," *Phys.Rev.* **D50** (1994) 5562–5579.

[6] A. Schofield and M. H. Seymour, "Jet vetoing and Herwig++," *JHEP* **1201** (2012) 078, `arXiv:1103.4811 [hep-ph]`.

[7] M. Bähr *et al.*, "Herwig++ Physics and Manual," *Eur. Phys. J.* **C58** (2008) 639–707, `arXiv:0803.0883 [hep-ph]`.

[8] S. Gieseke *et al.*, "Herwig++ 2.5 Release Note," `arXiv:1102.1672 [hep-ph]`.

[9] K. Arnold, L. d'Errico, S. Gieseke, D. Grellscheid, K. Hamilton, *et al.*, "Herwig++ 2.6 Release Note," `arXiv:1205.4902 [hep-ph]`.

[10] **ATLAS Collaboration** Collaboration, G. Aad *et al.*, "Measurement of dijet production with a veto on additional central jet activity in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector," *JHEP* **1109** (2011) 053, `arXiv:1107.1641 [hep-ex]`.

[11] **ATLAS Collaboration** Collaboration, G. Aad *et al.*, "Measurement of Dijet Azimuthal Decorrelations in pp Collisions at $\sqrt{s} = 7$ TeV," *Phys.Rev.Lett.* **106** (2011) 172002, `arXiv:1102.2696 [hep-ex]`.

[12] **CMS Collaboration** Collaboration, V. Khachatryan *et al.*, "Dijet Azimuthal Decorrelations in pp Collisions at $\sqrt{s} = 7$ TeV," *Phys.Rev.Lett.* **106** (2011) 122003, `arXiv:1101.5029 [hep-ex]`.

[13] **CMS Collaboration** Collaboration, S. Chatrchyan *et al.*, "Measurement of the Ratio of the 3-jet to 2-jet Cross Sections in pp Collisions at $\sqrt{s} = 7$ TeV," *Phys.Lett.* **B702** (2011) 336–354, arXiv:1106.0647 [hep-ex].

[14] **CMS Collaboration** Collaboration, S. Chatrchyan *et al.*, "Measurement of the Inclusive Jet Cross Section in pp Collisions at $\sqrt{s} = 7$ TeV," *Phys.Rev.Lett.* **107** (2011) 132001, arXiv:1106.0208 [hep-ex].

[15] **ATLAS Collaboration** Collaboration, G. Aad *et al.*, "Study of Jet Shapes in Inclusive Jet Production in pp Collisions at $\sqrt{s} = 7$ TeV using the ATLAS Detector," *Phys.Rev.* **D83** (2011) 052003, arXiv:1101.0070 [hep-ex].

[16] **ATLAS Collaboration** Collaboration, G. Aad *et al.*, "Properties of jets measured from tracks in proton-proton collisions at center-of-mass energy $\sqrt{s} = 7$ TeV with the ATLAS detector," *Phys.Rev.* **D84** (2011) 054001, arXiv:1107.3311 [hep-ex].

[17] **CMS Collaboration** Collaboration, V. Khachatryan *et al.*, "First Measurement of Hadronic Event Shapes in pp Collisions at $\sqrt{(s)} = 7$ TeV," *Phys.Lett.* **B699** (2011) 48–67, arXiv:1102.0068 [hep-ex].

[18] **ATLAS Collaboration** Collaboration, G. Aad *et al.*, "Measurement of multi-jet cross sections in proton-proton collisions at a 7 TeV center-of-mass energy," *Eur.Phys.J.* **C71** (2011) 1763, arXiv:1107.2092 [hep-ex].

[19] A. Buckley, J. Butterworth, L. Lonnblad, D. Grellscheid, H. Hoeth, *et al.*, "Rivet user manual," *Comput.Phys.Commun.* **184** (2013) 2803–2819, arXiv:1003.0694 [hep-ph].

[20] A. Martin, R. Roberts, W. Stirling, and R. Thorne, "NNLO global parton analysis," *Phys.Lett.* **B531** (2002) 216–224, arXiv:hep-ph/0201127 [hep-ph].

[21] **TASSO Collaboration** Collaboration, W. Braunschweig *et al.*, "Global jet properties at 14 GeV to 44 GeV center-of-mass energy in e^+e^- annihilation," *Z.Phys.* **C47** (1990) 187–198.

[22] **Particle Data Group** Collaboration, C. Amsler *et al.*, "Review of Particle Physics," *Phys.Lett.* **B667** (2008) 1–1340.

[23] **OPAL Collaboration** Collaboration, G. Abbiendi *et al.*, "Measurement of event shape distributions and moments in $e^+e^- \to$ hadrons at 91 GeV - 209 GeV and a determination of alpha(s)," *Eur.Phys.J.* **C40** (2005) 287–316, arXiv:hep-ex/0503051 [hep-ex].

[24] **OPAL Collaboration** Collaboration, G. Abbiendi *et al.*, "A Simultaneous measurement of the QCD color factors and the strong coupling," *Eur.Phys.J.* **C20** (2001) 601–615, arXiv:hep-ex/0101044 [hep-ex].

[25] **OPAL Collaboration** Collaboration, K. Ackerstaff *et al.*, "Measurements of flavor dependent fragmentation functions in $Z^0 \to q\bar{q}$ events," *Eur.Phys.J.* **C7** (1999) 369–381, arXiv:hep-ex/9807004 [hep-ex].

[26] **JADE collaboration, OPAL Collaboration** Collaboration, P. Pfeifenschneider *et al.*, "QCD analyses and determinations of α_s in e^+e^- annihilation at energies between 35 GeV and 189 GeV," *Eur.Phys.J.* **C17** (2000) 19–51, arXiv:hep-ex/0001055 [hep-ex].

[27] **JADE Collaboration** Collaboration, P. Movilla Fernandez, O. Biebel, S. Bethke, S. Kluth, and P. Pfeifenschneider, "A Study of

event shapes and determinations of alpha-s using data of e^+e^- annihilations at $\sqrt{s} = 22$ GeV to 44 GeV," *Eur.Phys.J.* **C1** (1998) 461–478, arXiv:hep-ex/9708034 [hep-ex].

[28] **DELPHI Collaboration** Collaboration, P. Abreu *et al.*, "Tuning and test of fragmentation models based on identified particles and precision event shape data," *Z.Phys.* **C73** (1996) 11–60.

[29] **DELPHI Collaboration** Collaboration, P. Abreu *et al.*, "Strange baryon production in Z hadronic decays," *Z.Phys.* **C67** (1995) 543–554.

[30] **ALEPH Collaboration** Collaboration, A. Heister *et al.*, "Studies of QCD at e^+e^- centre-of-mass energies between 91 GeV and 209 GeV," *Eur.Phys.J.* **C35** (2004) 457–486.

[31] **ALEPH Collaboration** Collaboration, R. Barate *et al.*, "Studies of quantum chromodynamics with the ALEPH detector," *Phys.Rept.* **294** (1998) 1–165.

[32] **ALEPH Collaboration** Collaboration, D. Buskulic *et al.*, "First measurement of the quark to photon fragmentation function," *Z.Phys.* **C69** (1996) 365–378.

[33] **ALEPH Collaboration** Collaboration, D. Decamp *et al.*, "Measurement of the charged particle multiplicity distribution in hadronic Z decays," *Phys.Lett.* **B273** (1991) 181–192.

[34] P. Nason, "A New method for combining NLO QCD with shower Monte Carlo algorithms," *JHEP* **0411** (2004) 040, arXiv:hep-ph/0409146 [hep-ph].

4 Treatment of Sub-Leading Colour in Event Generators

4.1 Introduction

We have seen in the previous sections that there are many ways in which we can improve the results of event generators. The implementation of missing physics should, at least after a tune, give improved results. The re-tune is a necessity as new physics will create new behaviour in the generator. This was discussed at length in the previous chapter. As outlined previously there are a number of sub-leading colour terms which are neglected in order to obtain reasonable algorithms for numerical implementation. Implementing some of these terms brings our generators more in accordance with what happens in nature and should also allow us to better predict experimental observables.

However, there are many sources of sub-leading colour in all of the parts of the event generator. The best way to proceed is to find which of these terms will have the largest impact on the predictions for experimental observables. Once the most important missing physics has been identified we can devise a set of new algorithms containing the desired physical behaviour. With the algorithm at hand we can then go to work producing parts of an event generator running on these new algorithms. This will be the first attempt to improve the colour description within the generator. In this chapter we will mainly consider the sub-leading colour missing from the parton shower as we expect that this will have the largest effect on any experimental observable.

4.2 Sources of Sub-Leading Colour

In order to improve the shower we must first identify the different sources of sub-leading colour and their relative impacts on experimental observables. We will begin with the treatment of colour within the eikonal approach, since it contains much of the colour structure used in the current parton showers. Once that has been considered we will then look into other possible sources of missing colour.

4.2.1 Colour in the eikonal approach

In the first chapter we were able to form a model of colour where the initial kinematics and the colour evolution were decoupled into separate matrices. The kinematic structure was contained in the hard matrix H and the colour structure contained in the colour metric S and the anomalous dimension Γ. The general forms of these matrices are

$$H = \quad H_d + H_o, \tag{4.1}$$

$$S = \quad S_d + S_o, \tag{4.2}$$

$$\Gamma = \quad \Gamma_d + \Gamma_o + i\Gamma_d^i + i\Gamma_o^i, \tag{4.3}$$

where the subscripts d and o represent on and off diagonal respectively and the superscript i indicates the imaginary part. In this section we will focus on the application of the eikonal model to gaps between jets phenomenology and consider only the colour flow basis, as this is the most important for numerical implementation.

4.2.1.1 The hard matrix H

For QCD $2 \to 2$ interactions at leading order all of the hard matrix elements will be proportional to poles in the s, t, or u channels. Events of interest will contain high s values as these are the highest energy

events which will produce the most radiation. The majority of events will therefore be those with low t or u or, equivalently, small angle scatterings.

In order to have a simple evolution algorithm which gives the correct hard cross section, the hard matrix is redefined in the way described in Chapter 1

$$\bar{H}_{ij} = \delta_{ij}(S_{ii}H_{ii} + \sum_{k\neq i} S_{ik}H_{ik}). \tag{4.4}$$

Recall that the evolution in the eikonal approach has the form

$$HS \to He^{-\xi\Gamma^\dagger}Se^{-\xi\Gamma}. \tag{4.5}$$

By modifying the hard matrix in the way suggested we have introduced errors in the full colour evolution. Each of the colour flows is evolved by a different factor than in the analytical calculation.

4.2.1.2 The soft metric S

The soft metric contains colour factors which arise from the forming of colour lines from an amplitude and a conjugate. The on-diagonal factors are when the amplitude and conjugate are the same state and the off-diagonal are when the states differ. The normalization of the states are typically chosen such that the on-diagonal factors are all unity. With this choice of normalization we find that the leading colour order of these terms are

$$S_d = 1, \tag{4.6}$$

$$S_o = \mathcal{O}(N_c^{-1}). \tag{4.7}$$

4.2.1.3 The on-diagonal real anomalous dimension Γ_d

The real on-diagonal elements of the anomalous dimension are those which are responsible for the reduction in gap cross section as the number of

emissions increase. They arise from the miscancellation between the real emission and virtual exchange due to the rapidity gap. The effects of these terms are included in the parton shower. However, there are sub-leading terms in the coefficient of these elements which are not included. There is one source of sub-leading colour which is included with these terms, and that is the different colour charges used for quarks and gluons. The different colour charges are included in order to ensure that the correct coefficient for the double logarithm is generated. The leading terms of the on-diagonal anomalous dimension are

$$\Gamma_d = \mathcal{O}(N_c). \tag{4.8}$$

4.2.1.4 The off-diagonal real anomalous dimension Γ_o

The off-diagonal elements of the anomalous dimension change one colour state into another and suppress the gap cross section. They can arise from both real emissions and virtual colour exchanges. These are not included in the parton shower. The leading terms of the off-diagonal anomalous dimension are

$$\Gamma_o = \mathcal{O}(1). \tag{4.9}$$

Previous investigations using the analytical model of the parton shower detailed in Ref. [1] have shown that these terms have minimal effects on the overall cross section.

4.2.1.5 The imaginary parts of the anomalous dimension Γ^i

The imaginary parts of the anomalous dimension are responsible for changing the colour state of either the amplitude or conjugate without a suppression to the cross section. They arise from virtual exchanges of colour between two initial state or two final state hard partons. The on-diagonal imaginary parts do not change the colour structure at all and

have no physical consequence. This can be understood as the fact that the redefinition of the anomalous dimension

$$\Gamma_{ij} \to \bar{\Gamma}_{ij} = \Gamma_{ij} + i\beta\delta_{ij} \tag{4.10}$$

does not change the cross section, where β is any real number. Neither of these two contributions to the anomalous dimension are included in the parton shower. The leading terms of the imaginary anomalous dimension are

$$\Gamma^i_o = \mathcal{O}(1). \tag{4.11}$$

Note that since the on-diagonal imaginary terms have no physical consequence their power of N_c is irrelevant.

4.2.2 Other sources of Sub-Leading Colour

While the eikonal approach is useful for treating colour when the emissions are soft, it does not contain the full picture with regards to colour. In the eikonal approach it is assumed that any additional emissions do not affect the kinematics of the hard process. We assume that the scale of each emission is negligible compared to that of any previous emission or interaction. Any emission of a gluon prior to the latest emission will be viewed as an indistinguishable part of the hard process. From this, in our eikonal approach we have assumed that all gluons are actually emitted from the initial hard partons. In real interactions all of the partons emitted at all scales will affect the colour evolution. They will be able to form dipoles which emit in addition to the initial hard partons. This is one of the differences between the parton shower and the eikonal model. In the parton shower all particles are evolved, rather than just the initial partons.

When we go beyond the eikonal model we need to consider the higher

dimensionality of colour space that is required. There are a set number of unique ways to connect the colour lines of four partons. As we increase the number of partons involved in the hard process then the number of colour structures present grows rapidly.

4.2.3 Relative sizes of effects

A naive approach to estimating the phenomenological impact of the different sub-leading colour effects would be to look at the power of N_c as stated above. However as we have stated above the actual effects of the different terms are different. From previous work we have found that the real off-diagonal terms have very little impact on the overall result. The imaginary off-diagonal terms are however very important. This is because they allow the transition between the colour structures which are available in the hard process and the colour singlet states. These colour singlets states present the possibility of very low levels of radiation occuring in an event. This will be very important, especially when it comes to the study of gaps between jets. One could say that the effect of changing the colour structure is important enough that it counteracts the lower colour prefactor. In addition to these changes the imaginary terms are always multiplied by a factor of π, which is able to counter a power of N_c.

4.3 Other sub-leading colour implementations

4.3.1 Plätzer-Sjödahl Approach

Plätzer and Sjödahl have implemented emission from sub-leading colour dipoles in a dipole shower, as detailed in Refs. [2]. In this approach they attempt to include emission from dipoles which are not colour connected in the leading colour structure. We will henceforth refer to this approach as the Plätzer-Sjödahl (PS) approach.

In the PS approach the probability of emitting from a dipole at the next step in the evolution $P_{i,j}$ for a dipole $\{i, j\}$ is proportional to

$$P_{i,j}(p_t^2, z; p_i, p_j) = K_{i,j}(p_t^2, z; p_i, p_j) C_{i,j}, \tag{4.12}$$

where $K_{i,j}$ is a kinematical factor and $C_{i,j}$ contains the colour structure of the splitting. The full colour factor is defined as

$$C_{i,j} = \frac{-1}{T_i^2} \frac{\langle M_n | T_i \cdot T_j | M_n \rangle}{\langle M_n | M_n \rangle}, \tag{4.13}$$

where $|M_n\rangle$ are the basis states for n partons and T_i is the colour generator for parton i. Physically this can be thought of as an emission from parton j in the amplitude and parton i in the conjugate amplitude. The basis used in the PS approach is that of the colour flow basis. In the leading colour limit this factor becomes

$$C_{i,j}^{LC} = \frac{1}{1 + \omega_i} \beta_{i,j} \tag{4.14}$$

where ω_i is 1 if i is a gluon or zero otherwise and $\beta_{i,j}$ is zero unless the partons i and j are colour-connected. In this limit the PS approach reproduces the results of the traditional dipole shower.

The PS approach however is capable of going beyond the leading colour limit. At each point in the colour evolution a colour basis for the current set of partons can be constructed and the colour factors $C_{i,j}$ can be determined exactly. This allows the colour factors of dipoles with sub-leading colour factors, which are neglected in the traditional dipole shower, to be calculated.

The PS dipole shower works in the same way as a traditional dipole shower but allows the possibility that dipoles with sub-leading colour could be the next ones that emit. By calculating the transverse momentum of

the next splitting for all possible dipoles the sub-leading colour effects of the real anomalous dimension can be included.

In comparison to the eikonal approach above, the PS approach focuses on generating the full colour factors in the real part of the anomalous dimension. The PS approach also contains colour effects not present in the eikonal approach. Since the full colour basis is calculated for the exact number of partons at each step the effects of the additional partons in the basis are included. This approach does not include the effects contained in the imaginary parts of the anomalous dimension.

The PS approach is currently limited in that it is only able to be applied to $e^+e^- \to q\bar{q}$ events. It has been implemented in a numerical code and results have been obtained for LEP observables. The numerical code was used to show that the results for most LEP observables were only modified by at most a few percent [2].

4.3.2 Nagy-Soper Approach

Nagy and Soper have generated a significant amount of analytical progress on sub-leading colour in showers, as shown in Refs. [3–6]. We will hereafter refer to this approach as the Nagy-Soper (NS) approach.

In this section we will focus on the latest results [6] in which a viable algorithm for going beyond the leading colour limit has been presented. The NS approach works by considering colour amplitudes as being composed of two types of objects, open strings and closed strings. The most basic open string is a colour connection between a quark and an anti-quark. Here, by anti-quark, we actually mean an anti-quark in the same temporal state (initial or final) or a quark in the opposite temporal state. Complexity can be added to the open string by attaching gluons to the colour line. A closed string is an open string in which the two ends are attached.

It is important to note that this string formalism is identical to the colour flow basis. In Chapter 1 the colour flow basis for $qg \to qg$ process is described in terms of open and closed strings and the colour flow basis for the $gg \to gg$ process is described in terms of closed strings[5].

One of the advantages of this colour string formalism is that $g \to gg$ splittings can always be expressed as two simpler $q \to qg$ splittings. As a result of this any splitting will just result in the insertion of gluons on strings and the creation or combination of strings.

The NS approach works in what they call the $LC+$ approximation. In this approximation they are able to generate the full colour factors for the soft and collinear and the collinear logarithms. Since they do not include the full effects of virtual colour exchanges they are unable to generate the correct colour factor for the soft logarithms.

The $LC+$ approximation is designed to be applied to a dipole shower. This approximation is defined as follows. We consider the emission of a gluon k from the dipole $\{i, j\}$. The colour generators are modified to be

$$T_i|M_n\rangle \to C(j,k)T_i|M_n\rangle. \tag{4.15}$$

The additional factor $C(j,k)$ is taken to be one only if the partons j and k are colour connected in the state $|M_{n+1}\rangle$ and zero otherwise. Here, we defined partons j and k to be colour connected if they are adjacent on a colour string.

While the NS approach is able to keep sub-leading colour terms it comes at a price. Each event has a weight factor associated with it. This weight is changed each time there is a splitting and depends on the colour factors involved in said splitting. It is also possible that these weights can be negative.

[5]The singlet terms in both cases can be expressed in terms of closed strings. Rather than just attaching the initial and final state gluons together, a quark loop is placed between.

While there is a discussion of the imaginary terms of the anomalous dimension there is no algorithm presented which can treat a changing colour structure during the evolution of the parton shower.

It is important to note that there is currently no code available that shows the results of the NS algorithms. However, a code based on these algorithms is said to be currently in development [6].

4.4 Colour Exchanges: General Consideration

In order to implement colour exchanges in an event generator we need to allow the colour structure to change. Additionally, in order to have a proper implementation, we need to have separate amplitude and conjugate colour structures. Allowing both of these conditions greatly increases the complexity of the algorithm.

Rather than work with the established generators like Herwig++ [7–9], we instead start with a stand-alone proof of concept in order to isolate the effects of different changes to the event generator. This new event generator will operate a hybrid shower, due to the advantages this shower methodology has over the parton shower.

4.4.1 Competing processes

In the original parton shower there is only one possible action that can happen at any given scale t. A parton can either split at the scale t or continue unchanged. When we add in colour exchanges we also have the possibility of one of the two colour structures changing at a scale t. Borrowing from the eikonal analytical resummation model we can view the colour exchange as being a competing process to a parton splitting. The scales for the different processes are calculated separately using their respective probabilities. The process with the highest scale is the one which is chosen to occur.

4.4.2 Concept of Evolution Time

The evolution in the parton shower depends on the concept of a time. This time is generally regarded as the time since the hard process occured. It is usually related to the virtuality. The higher virtuality interactions occur at earlier times, and the partons which are almost on-shell are evolved only at the very end of the shower. The actual implementation of time depends on the choice of the shower mechanism. There is a different behaviour for the parton shower relative to the hybrid and dipole showers.

4.4.2.1 Parton shower time evolution

In the current generation of event generators the evolution time is considered to be a property of each particle as it transitions through the shower. The highest evolution time is set by the colour connections in the hard process. Each evolution time is completely separate and each parton can be evolved separately, barring energy-momentum conservation concerns. In Herwig++ one of the partons in the hard process is chosen at random and then it and all its children are evolved until the minimum time is reached. Once this is done another hard parton is chosen until all of the partons in the process have been completely evolved. It is important to note that it is possible for evolution times of different partons evolved later to be greater than that of earlier ones owing to this random choice. An example of final state evolution demonstrating the different times in the parton shower is given in Figure 37.

In this particular example there are five different splittings at times denoted by t_i for the i-th splitting respectively. From the parton shower algorithm we are guaranteed that the times obey

$$t_1 \geq t_2, \tag{4.16}$$

$$t_3 \geq t_4 \geq t_5. \tag{4.17}$$

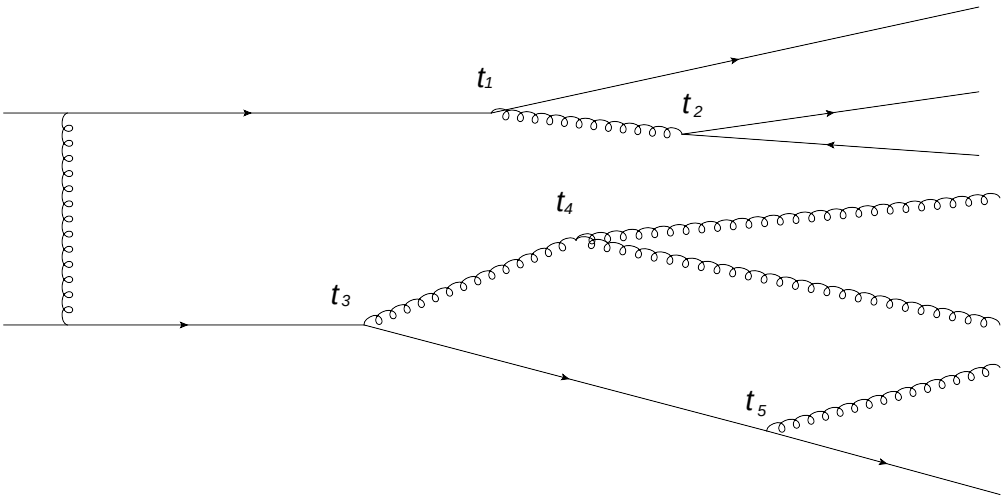

Figure 37: An example of a final state evolution in the parton shower. The symbols t_i represent the time of a given splitting i.

However, there is no guarantee that t_3 will be less than t_1. This means that interactions that occur at higher evolution times can be generated earlier than those at lower times, purely due to the random choice of the hard parton which is chosen to be evolved initially.

4.4.2.2 Dipole shower time evolution

In a dipole shower such as SHERPA [10], the highest time dipole is chosen to evolve first. Since energy and momentum are conserved at each interaction it is possible to switch between dipoles at any point in order to continue the evolution. Any dipoles which have times below the minimum are no longer evolved. An example of final state evolution demonstrating the different times in the dipole shower is given in Figure 38.

In this example two dipoles undergo four splittings at times t_1 to t_4. However in this case it is guaranteed that the evolution times obey

$$t_1 \geq t_2 \geq t_3 \geq t_4. \tag{4.18}$$

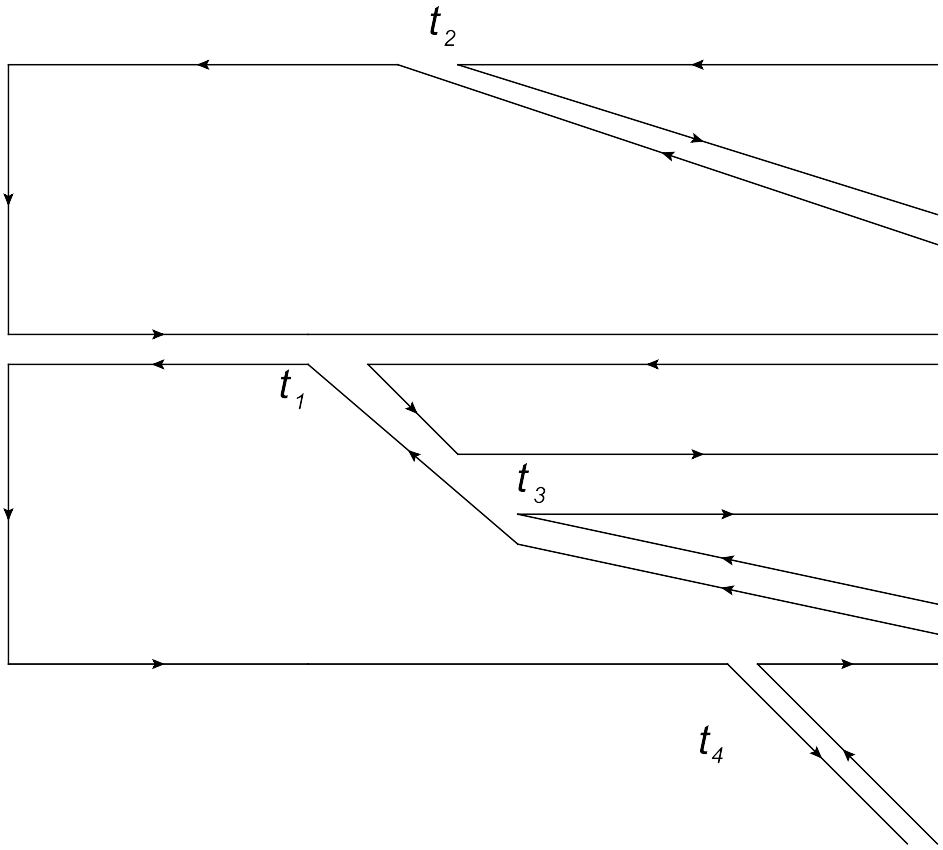

Figure 38: An example of a final state evolution. The symbols t_i represent the time of a given splitting i. Since the times of the two dipoles are interleaved the upper dipole can evolve before the lower one can be evolved a second time.

4.4.2.3 Hybrid time evolution

In the hybrid approach, used in PYTHIA 8 [11], each of the partons has their own shower time as in the parton shower. However, the kinematics are reconstructed after each splitting in a method similar to the dipole method. This allows any of the partons to be chosen to evolve at any given time. The choice of PYTHIA 8 is that the final state evolution is performed first and then any initial evolution is done.

In the hybrid shower we will also get splittings like that shown in Figure 37. However in this case it is possible to always evolve the parton which has the lowest evolution time due to the conservation of energy

and momentum.

4.4.2.4 Challenges

When colour exchanges are added, there exist interactions which can change the structure of the whole event, and hence modify the times of all partons. Thus it no longer makes sense to consider each parton as being able to be evolved separately. Instead, in our implementation only the highest time parton can evolve. In order to implement a universal time system we need to use either a hybrid or dipole shower. For the stand-alone program we choose to use the hybrid shower, as this is closer to Herwig++ than the dipole shower.

To correctly implement colour exchanges we must allow the amplitude and conjugate to have different colour structures. This results in differing scales depending on which of the two is chosen in a given stage of the evolution. Any algorithm aiming to fully implement these colour changes must take the effects of this change into account.

4.4.3 Variable colour structure

In a traditional parton shower, the colour structure is viewed as a solid frame to which additional radiation is attached. In terms of colour structure this is exactly what happens in the leading colour limit of the eikonal picture. To improve the parton shower implementation we can begin to borrow additional properties from this picture. Rather than attempting to implement all of the sources of sub-leading colour at once, we will begin with minimal changes.

The first change is to allow the colour frame to change. We define the colour frame as the colour structure of the hard process. Later on we will refer to these as junctions. To get the correct physical picture we will need to allow the colour frames of the amplitude and conjugate to

differ. This is because any exchange of colour will only change the colour structure of either the amplitude or the conjugate and not both.

If the colour frames of the amplitude and conjugate differ then it is possible for partons to have a different scale in each. There is no exact physical reasoning for when to choose between the amplitude and conjugate. We make the choice that at the beginning of each loop of the shower we should choose one of the two randomly. Once the amplitude or conjugate is chosen, it is retained throughout that iteration of the loop. Any colour exchanges will only be able to change the chosen frame.

4.5 Colour Exchanges: Phenomenological Model

In this section we will consider a phenomenological model of colour exchanges for a parton shower. The goal of this model is attempt to improve the parton shower by implementing variable colour structures. The full QCD behaviour is far too complex to try to implement all at once. Instead, we will consider a much simpler model where the colour structure of the hard process is mostly separated from the evolution. By doing this we can leave most of the structure of the parton shower intact while adding in new physics.

This model has been adapted into a numerical algorithm and implemented in a program as a stand-alone parton shower. However, this program is still not at a stage where meaningful results can be produced.

4.5.1 Overview

In this section we outline the generator behaviour as a whole, before more details on each of the parts is given in the subsequent sections. The flow of the overall program is shown in Figure 39. The algorithm acts on two main objects. These objects are `particle`s and `event`*s*. All of the physical behaviour is modelled through the interactions between these

two objects.

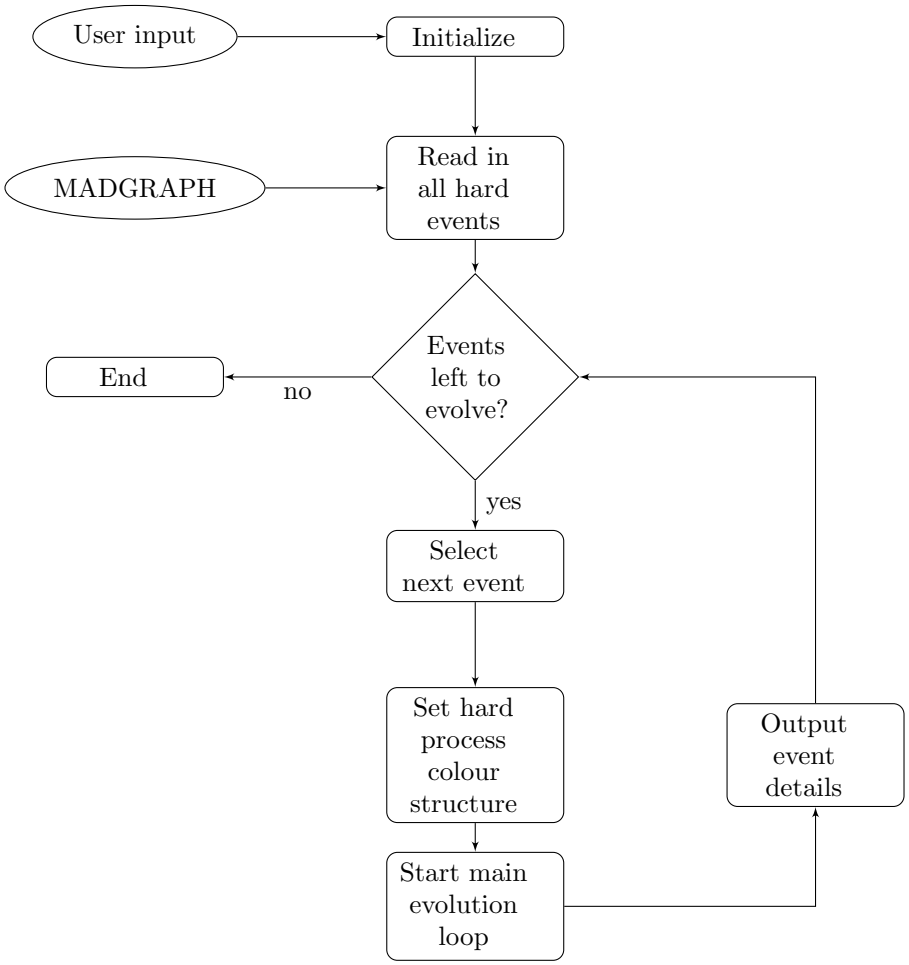

Figure 39: The overall flow of the algorithm.

The `particle` object contains the properties associated with a physical particle. All particles i will have a set of momenta p_i chosen in a particular common frame. All of the `particle`s within the algorithm are partons and therefore have colour lines $\{c_i, \bar{c}_i\}$ associated with them. These colour lines connect them to other partons. Each parton will have one or two partners. The maximum scales t of these `particle`s are given by the virtuality associated with a particle i and its partner j

$$t = 2p_i \cdot p_j, \tag{4.19}$$

where we have assumed that all partons are massless. The `particle`s also carry a scale ratio R_i, which is the ratio of their current interaction scale to their maximum possible interaction scale.

An `event` is a container for one specific high energy collision. Initially it will contain the set of `particle`s involved in the hard interaction. After every splitting more `particle`s will be added to the `event`. The `event` also contains the current colour structures of the amplitude and conjugate amplitude. By doing this we can separate any complicated colour changes from the `particle`s and consider them independently.

Within the algorithm we consider the colour structures to be a set of connections that map one colour line to another within the hard process. Different colour structures connect the colour lines in the hard process in different ways. We will call one specific configuration of connections a junction. The junction $J(i, C_j)$ maps the colour line i to the colour line k to which it is connected in the hard process through the colour structure C_j. The junction is the map

$$J(i, C_j) = k. \tag{4.20}$$

An example of this map for one of the colour structures of $gg \to gg$ is shown in Figure 40. The full list of junction values is given in Table 5. While in this example the junction is only connecting initial colour lines to final ones, it is possible that colour lines from the same temporal state can be connected together as well.

i	1	2	3	4	5	6	7	8
$J(i, C)$	5	8	7	6	1	4	3	2

Table 5: An example of a junction values for $gg \to gg$.

The behaviour of the algorithm can be customized with a set of inputs at run-time. These inputs control whether colour exchanges can occur,

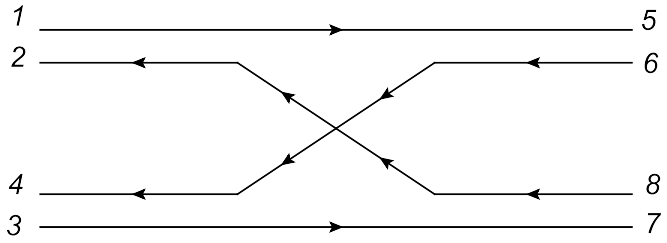

Figure 40: An example of a junction for $gg \to gg$.

whether to use a separate amplitude and conjugate amplitude, as well as many other properties.

4.5.2 Input stage

The program itself contains no methods for generating a realistic hard process. Instead, events need to be generated first with MADGRAPH [12]. These are considered as a set of momenta and flavours for the incoming and outgoing particles. It is also possible for the colour structures to be read directly from MADGRAPH. This is left as an option to the end user.

4.5.3 Initial Colour Selection

If the colour is not read in from MADGRAPH then it is chosen at this point. The choice of colour structure is obtained by using kinematical weights. At this point we consider all processes to have the same colour structure for both the amplitude and the conjugate amplitude. In order to generate the correct cross section we make the approximations that were outlined in Chapter 1.

In the hard process there are N colour lines. The value of N depends on the process involved. Each quark or anti-quark has one colour line, while each gluon has two. The partons in the hard process are assigned colour line values from 1 to N.

4.5.4 The Main Loop

Once an event has been selected it must now be evolved with the parton shower. The flow of the parton shower is shown in Figure 41. The parton shower is composed of multiple runs of the main loop. Each run through the main loop contains a choice of amplitude and the possibility of either a splitting or a colour exchange.

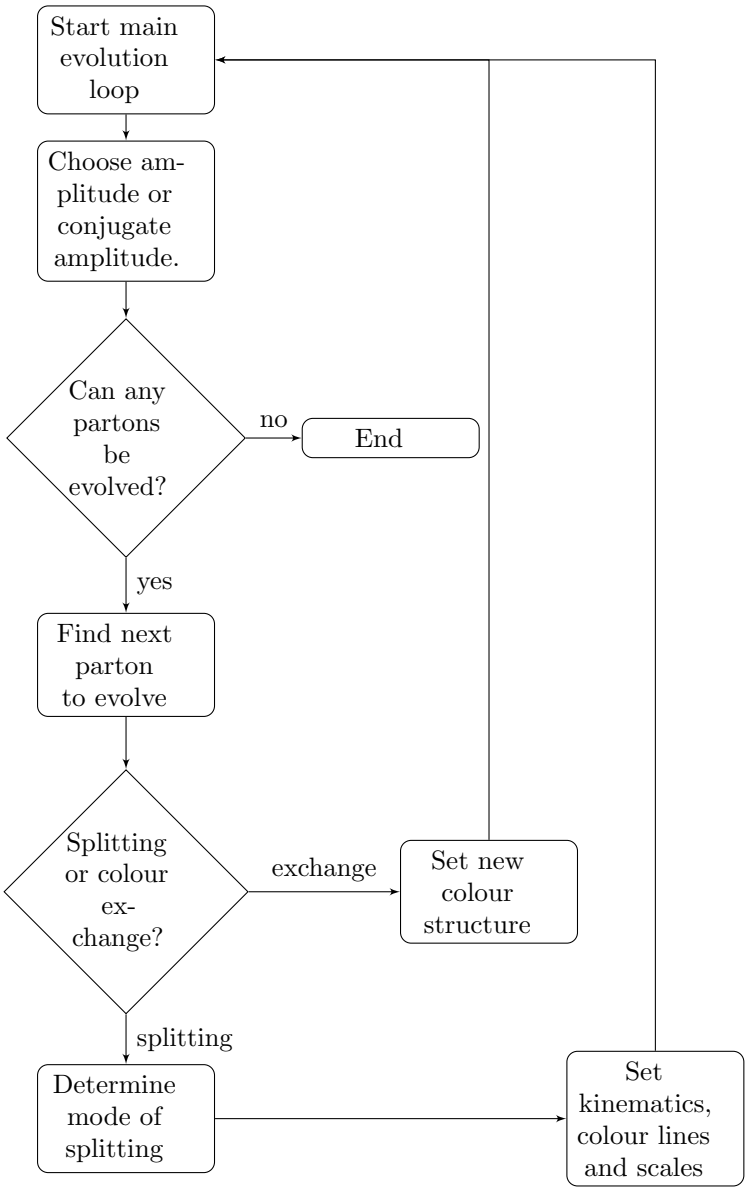

Figure 41: The main event loop for the new parton shower algorithm.

4.5.5 Choice of Amplitude

At the start of the loop the algorithm will make a choice of whether to use the colour structure in the amplitude or the colour structure in the conjugate amplitude. With our current implementation the colour structures will be the same at the start of the parton shower. However, after colour exchanges, these may differ. This means that particles may have different scales depending on which of these two amplitudes are chosen.

There is no restriction on how the algorithm makes the choice of using the amplitude or conjugate amplitude. The probability of choosing either one of those can be equal, or it can be weighted by the kinematics of the hard process if desired. The current implementation chooses either the amplitude or conjugate amplitude with equal probability using a random number.

If the colour structure in the chosen amplitude differs from that of the previous run through the loop then the scales t of all partons connected to the hard process need to be changed. The partons which are connected to the hard process are those which have colour lines with values less than or equal to N. If we define the scale of a particle in colour structure C_i to be t_i, then we choose the scale t_j in colour structure C_j to be

$$t_j = \frac{t_i}{t_i^m} t_j^m = R t_j^m, \tag{4.21}$$

where t_i^m and t_j^m are the maximum scales for the particle in colour structures i and j respectively. We preserve the ratio of the current scale to the maximum scale when the colour structure changes. This is the default behaviour. We also include another possible behaviour that

restricts the scale to only decrease under a colour structure change

$$t_j = \min\left(\frac{t_i}{t_i^m} t_j^m, t_i\right). \tag{4.22}$$

The choice of which of these two behaviours to use is left to the end user.

4.5.6 Choice of Interaction

Once the amplitude has been chosen the scales for each of the partons in the initial and final state are determined. The default choice of scale for a parton is the virtuality between said parton and its partners. A quark or anti-quark will only have one partner and therefore one scale per colour structure. Since gluons have two colour lines there is the possibility of being connected to two different partons, which results in two different scales per colour structure. The greater of the two scales sets the maximum possible scale of interaction.

The next step in the algorithm is to determine the next scale of interaction. This is done by checking what the next interaction scale is for each parton in both the initial and final states. For each parton there are two different kinds of interaction of which they can be involved in. The first is a splitting, which is where a parton will split in to two other partons. This is treated in the same way as any other parton shower algorithm which accounts for a recoiling parton. The scale of the next splitting t_s is given by solving

$$R_s = \exp\left(-\alpha_s^o \int_{t_s}^{t_r} \frac{dt}{t} \int_{z-}^{z+} dz P^o(z)\right), \tag{4.23}$$

where t_r is the scale of the parton prior to this stage of evolution, α_s^o is a constant overestimate depending on the event in question, $P^o(z)$ are overestimates of the DGLAP splitting kernels [13–15] and R_s is a random number in the range [0,1]. This expression is solved with an analytical

rearrangement similar to that described in Chapter 1.

The second method of interaction is a colour exchange, where a virtual gluon is exchanged between two partons. We make the simplifying choice that only partons which have colour lines connected to the hard process can make these exchanges. The scale of the next colour exchange t_e is given by solving

$$R_e = \exp\left(-\lambda\alpha_s^o \int_{t_e}^{t_r} \frac{dt}{t} \sum_j |\mathrm{Im}(\Gamma_{ji})|\right), \tag{4.24}$$

where Γ_{ji} is the anomalous dimension element, i is the index of the current colour structure of the chosen amplitude, λ is a colour strength correction factor and R_e is a random number in the range [0,1]. The values of the anomalous dimension are not positive-definite. In order to obtain a reasonable probability for colour exchanges we have taken the absolute value of the elements instead.

The colour strength factor λ requires further discussion. Our current approach only allows partons connected to the hard process to exchange colour through virtual gluons. A more realistic approach would allow colour exchanges between any of the partons involved in the evolution. This is far too complex for implementation within our current algorithm. Instead we add the colour strength factor λ which is meant to take in to account the effects of such additional exchanges. We treat λ as a parameter which is not fixed, but variable in order to match experimental observables.

Since we have made a number of assumptions in order to have a reasonable algorithm this is only a very crude model of colour exchanges. A better algorithm might take into account weights for events from the negative values of the anomalous dimensions, or allow colour exchanges between all partons.

The next interaction scale of the parton i, t_i, is now defined as the

greater of the two possible scales t_s and t_e

$$t_i = \max(t_s, t_e). \tag{4.25}$$

The parton with the largest interaction scale is evolved next. The method of interaction is given by that which provided the larger scale.

In subsequent passes through this section there will be a mix of partons which already have their next scales set, as well as new partons or partons which were involved in the previous splitting or exchange. The partons which don't have their next scale set use the above methodology to gain their next scale. The parton with the greatest scale is then chosen to evolve.

4.5.7 Splitting

If a parton p is chosen to split then it will behave as in the normal hybrid shower. With our hybrid shower we consider the splitting of a parton p with a colour connected recoiler parton r,

$$p + r \to p' + q + r'. \tag{4.26}$$

The first step is to determine the type of splitting which occurs. This determines both the flavours and the colours of the children p' and q produced by the splitting.

For a final-state quark there is only one possible splitting, which is where a quark will emit a gluon. We will however treat this as two separate splittings, one where the gluon takes the majority of the momenta of the parent quark and one where the gluon instead takes the majority. These two cases only differ by how their scale is set.

For an initial-state quark there are two possibilities for splitting. Since we are considering backwards evolution, the quark that is splitting was

either generated by a gluon or another quark. These two splittings have different colour assignments and scales, as well as PDF factors.

For a final-state gluon there are again two possibilities. The gluon can either split in to two gluons or it can produce a quark-antiquark pair. The colour assignment for the quark-antiquark production mode is trivial. For the two gluons there are two possible colour assignments. However, only one of these colour assignments will be possible as the colour line which connects the parton with the recoiler must be the one to split.

For an initial-state gluon the two possibilities are the same as the quark, but reversed. The gluon could have either been produced through emission from a quark or through the splitting of another gluon. Again, these different combinations will have different colour and PDF factors associated.

The weight W_i of a particular splitting i can be defined as

$$W_i = \frac{P_i}{\sum_j P_j} \tag{4.27}$$

where the factors P_i contain an integration over the splitting kernels and PDFs. The mode of splitting is chosen randomly from all possible modes using these weights.

Once the mode of splitting has been decided the kinematics must then be constructed. The kinematics for this process are independent of the flavours of the partons. The kinematics only depend on the temporal states of the splitting parton and the recoiler. One particular choice of parameterization for the kinematics is

$$p' = \alpha p + \beta r + \gamma k_t, \tag{4.28}$$

$$q = \chi p + \delta r + \omega k_t, \tag{4.29}$$

$$r' = \xi r. \tag{4.30}$$

The values of the parameters for the different temporal states are given in Table 6. Here v is the virtuality, which is the same as the scale in the default implementation, and s is defined as the maximum virtuality of the splitting parton and recoiler pair

$$s = 2p \cdot r. \tag{4.31}$$

State	α	β	γ	χ	δ	ω	ξ
FS/FS	z	$(1-z)v/s$	1	$(1-z)$	zv/s	-1	$(1-v/s)$
FS/IS	z	$(1-z)v/s$	1	$(1-z)$	zv/s	1	$(1+v/s)$
IS/FS	$1/z$	$(1-z)v/s$	$1/z$	$1/z-1$	zv/s	$1/z$	$(1+v/s)$
IS/IS	$1/z$	$(1-z)v/s$	$1/z$	$1/z-1$	zv/s	$-1/z$	$(1-v/s)$

Table 6: Kinematic properties for splittings with the four possible temporal state combinations.

In order to have the correct splitting probabilities while using the simplified splitting kernels we must apply vetoes. There will be vetoes for α_s, the kernel overestimates and, if we have an initial-state splitting, the PDFs. We must also use an additional veto for angular ordering as we are using virtuality as our evolution time. If the splitting is vetoed then the parton is set to evolve again, but with a maximum scale equal to the scale of the splitting.

The colour line assignment during splitting is important. This is because the connections of the colour lines determine scales. For quark splittings there is only one possible assignment of colour lines. For gluon splittings there are however two possible ways of assigning the colour lines. The colour lines are assigned in Table. 7.

Once the splitting has occured, the scales of all partons evolved must be changed. This does not only include the parton that split, but also the recoiling parton and all of the partons which those two are colour connected to. The scales of the children $t_{i,j}$ are related to that of the

Splitting Type	c_q	$\bar{c}_q$	$c_{\bar{q}}$	$\bar{c}_{\bar{q}}$	c_g^1	$\bar{c}_g^1$	c_g^2	$\bar{c}_g^2$
FS $q \to qg$	n	-	-	-	c_p	n	-	-
IS $q \to qg$	n	-	-	-	n	c_p	-	-
FS $q \to gq$	n	-	-	-	c_p	n	-	-
IS $q \to gq$	-	-	-	n	c_p	n	-	-
FS $g \to q\bar{q}$	c_p	-	-	$\bar{c}_p$	-	-	-	-
IS $g \to q\bar{q}$	c_p	-	$\bar{c}_p$	-	-	-	-	-
FS $g \to \bar{q}q$	c_p	-	-	$\bar{c}_p$	-	-	-	-
IS $g \to \bar{q}q$	-	c_p	-	$\bar{c}_p$	-	-	-	-
FS $g \to gg$	-	-	-	-	c_p	n	n	$\bar{c}_p$
FS $g \to gg$	-	-	-	-	n	$\bar{c}_p$	c_p	n
IS $g \to gg$	-	-	-	-	c_p	n	$\bar{c}_p$	n
IS $g \to gg$	-	-	-	-	n	$\bar{c}_p$	n	c_p

Table 7: All possible colour line assignments. c_i denotes the colour line of parton i while $\bar{c}_i$ denotes the anti-colour line of the same parton. If a new colour line is generated in the splitting it is referred to as n. For initial state $g \to q\bar{q}$ and $g \to \bar{q}q$ the first child is the initial state (anti-)quark and the second child is the final state (anti-quark).

parent t by

$$t_i = zt, \tag{4.32}$$

$$t_j = (1-z)t. \tag{4.33}$$

4.5.8 Colour Exchange

If a parton is chosen to produce a colour exchange then we must consider what happens to the colour structure. In our algorithm we have previously made a choice of amplitude or conjugate amplitude for the colour structure. Only the colour structure which was chosen can be modified by a colour exchange. The probability P_{ij} of transitioning to a state C_j from state C_i is given by

$$P_{ij} = \frac{|\mathrm{Im}(\Gamma_{ji})|}{\sum_k |\mathrm{Im}(\Gamma_{ki})|} \tag{4.34}$$

The new colour state is chosen from the potential set of colour states $\{C_j\}$ with weights given by the above probabilities. If the colour structure has changed then the connections between partons in the hard process are now different. This is shown in Figure 42. This will result in different scales for partons which are connected to the hard process through their colour lines. The partons which are not connected to the hard process will be left unchanged.

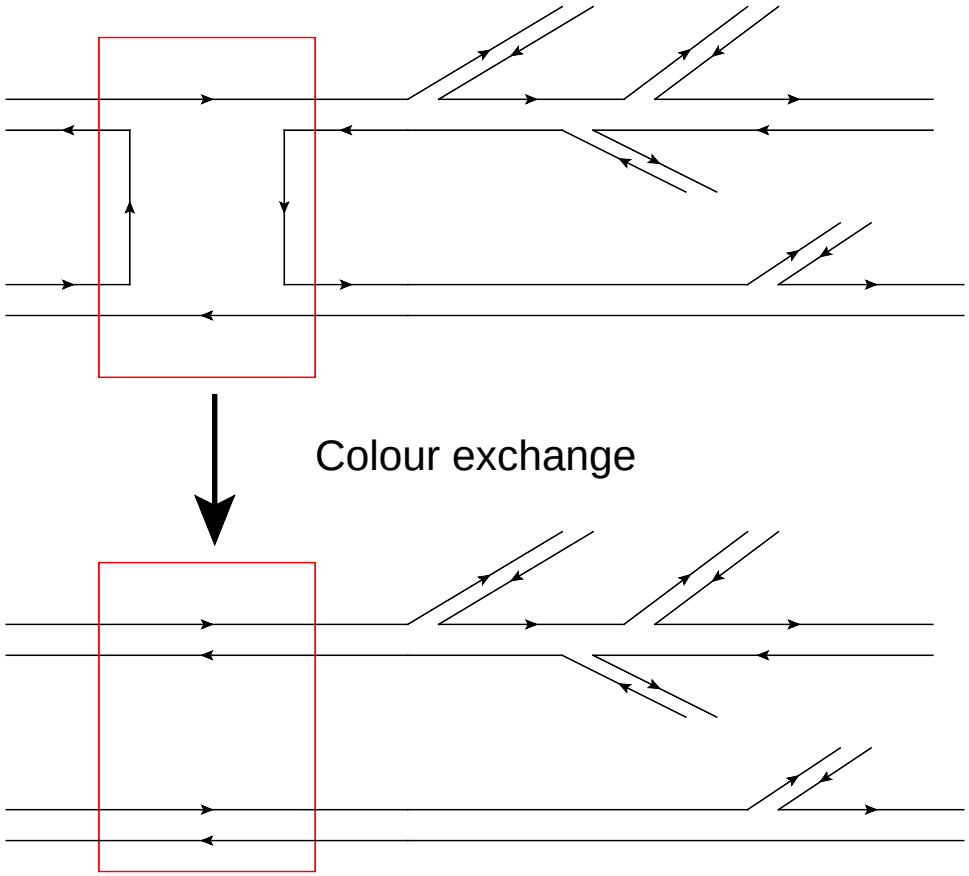

Figure 42: A colour exchange resulting in the change of a junction. The rest of the event is unchanged.

4.5.9 End of the loop

Once a colour exchange, splitting or veto has occured the current loop of evolution has ended. If there are no more partons which can be evolved at this point then the shower will terminate.

4.5.10 End of the shower

The shower terminates when all possible interactions are below a shower cutoff value, which we define to be 1 GeV. At this point the event consists as a set of initial and final state partons with their respective momenta and flavours, as well as the colour structures of the amplitude and conjugate amplitude. Since we include no hadronization mechanism there is no problem if the final colour structures differ. However, if we were to include a hadronization mechanism we would a method for choosing one of the two colour structures. One possibility is simply to select one of the two colour structures at random. An alternative method would be to allow the shower to continue, but with only colour exchanges allowed. Once the two colour structures match, this extension of the shower would terminate.

The final events are not weighted in any way, regardless of whether the final colour structures of the amplitude and conjugate amplitude match. The output of the shower is handled by HepMC [16] and stored as data files.

4.5.11 Example event

In order to give a clearer explanation of the algorithm we will now consider one specific event that might occur. This is shown in Figure 43. In this example we consider $qq \to qq$ interactions, which only have two colour states, and restrict ourselves to just the evolution of the final state. Prior to the parton shower, denoted as loop pass 0, the colour structures of the amplitude and conjugate amplitude are the same and no evolution has yet occured. There are only two quarks in the final state.

The first stage of evolution from the parton shower is in loop pass 1. At the start of this loop the colour structure in the amplitude is chosen. The evolution mode of choice is a splitting, which in this case is $q \to qg$. The gluon is represented as a pair of colour and anti-colour lines. This

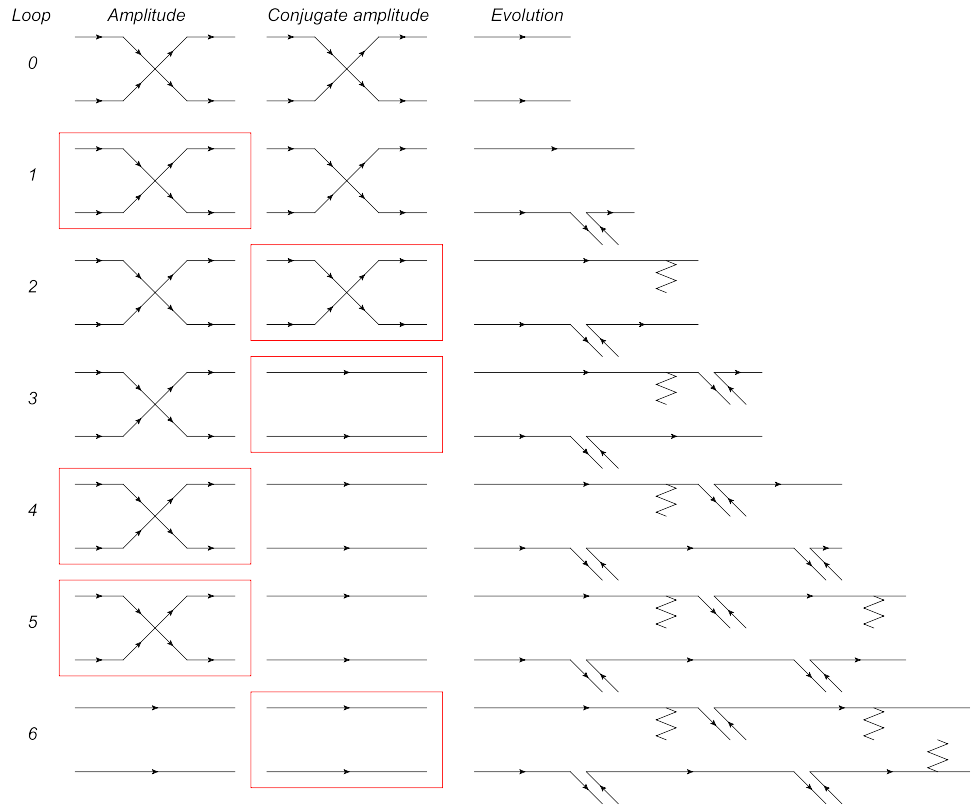

Figure 43: A simplified example of an event within the algorithm. The event shown here is for a $qq \to qq$ interaction where only the final-state partons are allowed to evolve. The red boxes denote which amplitude is being used within the evolution section. Zigzag lines are used to indicate colour exchanges. The evolution time grows with distance from left to right.

stage of the evolution ends with a two quarks and a gluon in the final state.

The next stage of evolution is in loop pass 2. This time the colour structure for the conjugate amplitude is chosen. The evolution choice this time is a colour exchange, which is denoted as a zigzag line on the diagram. The colour exchange may or may not result in a change of the colour structure of the conjugate amplitude. This stage of the evolution ends with two quarks and a gluon in the final state.

The third stage of evolution is in loop pass 3. Again, the colour

structure for the conjugate amplitude is chosen. However, we see that the previous colour exchange did in fact change the colour structure into the only other possible state, which is that of a colour singlet. This means that the scales for the partons connected to the hard process, which for the final state are the quark at the top and the gluon at the bottom, need to have their scales re-evaluated. Once these scales have been determined there is another $q \to qg$ interaction. This stage of the evolution ends with a two quarks and two gluons in the final state.

The subsequent loop passes repeat behaviour which has already been described in the previous passes. In loop pass 4, the colour structure of the amplitude is chosen and a $q \to qg$ interaction occurs. In loop pass 5, the colour structure of the amplitude is chosen again and a colour exchange occurs. Looking at loop pass 6, it is clear that this colour exchange has resulted in the move to a singlet state, this time for the amplitude. Finally, we have another colour exchange, this time using the colour structure of the conjugate amplitude.

This example shows the general behaviour of the algorithm. At each stage either the colour structure of either the amplitude or conjugate amplitude is chosen. A method for evolution, either a splitting or colour exchange, is then chosen. The splittings are stored within the evolution part and are separate from the colour structures stored in the amplitudes. The colour exchanges have a possibility of changing the colour structures within the amplitudes.

4.6 Implementation in regular parton showers

While the approach outlined above explains how we might implement basic colour exchanges in a stand-alone code, it will not work for implementation in the normal parton shower. This is because it is not possible to have the universal time, as outlined above. Instead we consider an alternative

algorithm here, which may give the same results.

Before the parton shower is run we will generate the colour exchanges. This is done by running the parton shower but with splittings replaced by colour exchanges. The scale at which each colour exchange occurs is recorded. Additionally, we will randomly choose whether to use the colour structure of the amplitude and conjugate amplitudes over a given time scale. How this choice is made, and the time scales involved, are left to the implementation of the algorithm. The simplest method would be to choose randomly between the colour structure of the amplitude and conjugate amplitude, as is done in the standalone algorithm. The time scale could be determined by viewing the changing of the amplitude and conjugate as competiting processes and determining which of the two will perform a colour exchange first. Once the colour structures for different scales have been determined the parton shower is run as usual, but with colour structures varying with scale as determined above.

This method is however not without flaws. Every time the colour structure changes, so do the scales involving the hard partons. As such, it would be possible to move to an earlier time than the colour change by changing the colour. Any such use of this model must address this flaw. This is solved in the standalone implementation by only allowing scales to decrease under a colour structure change.

Creating an implementation within existing generators will be an important step for phenomenology. While a proof of concept can outline the importance of including these effects, it will lack the full predictive power of the established event generators. Having two different implementations, one in the standalone shower and one in a traditional parton shower, allows us to cross-check the results of said implementations. While the two implementations may not give identical results, we might expect that they would give a similar trend for most observables. By studying these

trends we could check if our algorithms are correctly implemented within the numerical codes.

4.7 Conclusion

We have seen in this chapter that there are a large number of sub-leading colour corrections which are not present in any current event generator. While not all of these have a major impact on phenomenology, some may have great importance at the currently accessible and future energies provided by colliders. The impact of these sub-leading terms requires a great deal of future study. In this chapter we have made the first step towards a greater understanding of these sub-leading effects.

Though we have a set of viable algorithms for numerical implementation, there is not yet a fully working code that is able to demonstate their impact.

References

[1] A. Schofield and M. H. Seymour, "Jet vetoing and Herwig++," *JHEP* **1201** (2012) 078, arXiv:1103.4811 [hep-ph].

[2] S. Plätzer and M. Sjödahl, "Subleading-N_c improved Parton Showers," arXiv:1206.0180 [hep-ph].

[3] Z. Nagy and D. E. Soper, "Parton showers with quantum interference," *JHEP* **0709** (2007) 114, arXiv:0706.0017 [hep-ph].

[4] Z. Nagy and D. E. Soper, "Parton showers with quantum interference: Leading color, spin averaged," *JHEP* **0803** (2008) 030, arXiv:0801.1917 [hep-ph].

[5] Z. Nagy and D. E. Soper, "Parton showers with quantum interference: Leading color, with spin," *JHEP* **0807** (2008) 025, arXiv:0805.0216 [hep-ph].

[6] Z. Nagy and D. E. Soper, "Parton shower evolution with subleading color," *JHEP* **1206** (2012) 044, arXiv:1202.4496 [hep-ph].

[7] M. Bähr *et al.*, "Herwig++ Physics and Manual," *Eur. Phys. J.* **C58** (2008) 639–707, arXiv:0803.0883 [hep-ph].

[8] S. Gieseke *et al.*, "Herwig++ 2.5 Release Note," arXiv:1102.1672 [hep-ph].

[9] K. Arnold, L. d'Errico, S. Gieseke, D. Grellscheid, K. Hamilton, *et al.*, "Herwig++ 2.6 Release Note," arXiv:1205.4902 [hep-ph].

[10] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, *et al.*, "Event generation with SHERPA 1.1," *JHEP* **0902** (2009) 007, arXiv:0811.4622 [hep-ph].

[11] T. Sjöstrand, S. Mrenna, and P. Z. Skands, "A Brief Introduction to PYTHIA 8.1," *Comput.Phys.Commun.* **178** (2008) 852–867, arXiv:0710.3820 [hep-ph].

[12] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, "MadGraph 5 : Going Beyond," *JHEP* **1106** (2011) 128, arXiv:1106.0522 [hep-ph].

[13] V. N. Gribov and L. N. Lipatov, "Deep inelastic e p scattering in perturbation theory," *Sov. J. Nucl. Phys.* **15** (1972) 438–450.

[14] G. Altarelli and G. Parisi, "Asymptotic Freedom in Parton Language," *Nucl. Phys.* **B126** (1977) 298.

[15] Y. L. Dokshitzer, "Calculation of the Structure Functions for Deep Inelastic Scattering and e^+ e^- Annihilation by Perturbation Theory

in Quantum Chromodynamics," *Sov. Phys. JETP* **46** (1977) 641–653.

[16] M. Dobbs and J. B. Hansen, "The HepMC C++ Monte Carlo event record for High Energy Physics," *Comput.Phys.Commun.* **134** (2001) 41–46.

5 Final Conclusion

We have shown throughout this thesis that the treatment of colour, both in the analytical and numerical approaches to simulating scattering processes, is important for predicting experimental observables. Within Herwig++ the old approach to colour line choices for the splitting of gluons resulted in unphysical behaviour. By modifying this approach we were able to better match the analytical approach.

Having made the changes to the colour implementation of Herwig++, we then moved on to investigating how this had affected the optimal tune of the model parameters. We observed that the changes we had made opened up a possibility for a number of tunes to be used, but in general these did not depart significantly from the currently used tune when it came to using them to predict experimental observables.

We have also explained how the implementation of sub-leading colour effects may have an impact on the final state of the parton shower. The true impact of a full implementation of sub-leading colour effects is still unknown and requires further investigation.