MANCHESTER 1824
The University of Manchester

Simulating jet processes with `CVolver`

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

2025

Fernando Torre González
School of Physics and Astronomy
Theoretical Particle Physics Group

Contents

Word count: 37136

List of figures

List of tables

Abstract

This thesis presents the continued development and application of the amplitude-level parton shower Monte Carlo `CVolver`, which is capable of resumming soft-gluon logarithms at full colour. The framework is built on a soft evolution algorithm formulated in the colour flow basis, which allows for a systematic treatment of colour suppression throughout the evolution. We detail the most up-to-date implementation of `CVolver`, including the methods used to steer the evolution in order to minimise event weights and maximise efficiency in computing the most relevant contributions. `CVolver` supports two operational modes: a *dedicated mode*, optimised for efficiently resumming soft-gluon effects in jet veto cross sections, and an *event generator mode*, which functions as a general-purpose parton shower. Using the dedicated mode, we carry out a systematic and comprehensive study of subleading-colour corrections in perturbative QCD processes. Specifically, we analyse the processes $Z \to q\bar{q}$, $H \to gg$, $q\bar{q} \to q\bar{q}$, $qg \to qg$, and $gg \to gg$, finding significant subleading-colour effects at the 5–30% level, with even larger corrections arising from interference terms. Notably, we observe that leading-colour evolution provides an accurate approximation to the full-colour result in $2 \to 2$ t-channel gluon exchange processes with forward scattering kinematics. Finally, we use the event generator mode of `CVolver` to investigate the differential structure of soft radiation, and find pronounced subleading-colour effects at the 10% level in various angular regions.

Declaration of Authorship

I hereby confirm that 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, trademarks 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=24420`), in any relevant Thesis restriction declarations deposited in the University Library, The University Library's regulations (see `http://www.library.manchester.ac.uk/about/regulations/`) and in The University's policy on Presentation of Theses.

Acknowledgements

There are a lot of people I would like to thank for their guidance and support throughout this project. The two most important are Jeff Forshaw and Simon Plätzer. Without their passion and perseverance, I would probably have given up multiple times over. Our outings in Graz will remain some of my best memories of the past few years.

I also need to thank the computing support of the Particle Physics Groups at the universities of Manchester and Vienna. I would not be surprised if I have broken some sort of record for the amount of computational resources I used. Special thanks to Conor Fitzpatrick in Manchester and Helmut Moser in Vienna for accommodating me.

I am very lucky to have had many friends supporting me through this journey, and I cannot do them justice in just a few paragraphs.

I will start with the non-physicists: James, Laora and Becca. I cannot imagine what the past few years would have been like without you, and I hope you know how important you are to me.

I continue with the physicists: Marion and Viola. You both kept me sane and reminded me to have fun. Even now. “Just thank yourself and send it.”

I also want to acknowledge my friends whom I do not get to see nearly enough: César, Christian, Marta. I know that no matter how far away we are, you’ll always be there for me.

Lastly, my parents, Virginia and Antonio, and my sister, Silvia. Without your support, I would not be here, sitting in Manchester, writing the acknowledgements for my PhD thesis.

Introduction

The Standard Model of particle physics has withstood decades of experimental scrutiny and boasts an impressive track record of predictive success. It encapsulates our most comprehensive understanding of three of the four fundamental forces of nature: electromagnetism, the weak interaction, and the strong interaction, and some notable milestones include the prediction and subsequent discovery of the W and Z bosons [1, 2], the top quark [3, 4], and the crowning achievement of the Higgs boson discovery [5–8]. Despite these successes, the Standard Model remains incomplete. It does not incorporate gravity, offers no explanation for the observed matter–antimatter asymmetry in the universe, and fails to account for dark matter and dark energy, which dominate the energy content of the universe. These unresolved questions, and more, motivate the ongoing search for physics beyond the Standard Model.

Among the interactions described by the Standard Model, the strong force—governed by Quantum Chromodynamics (QCD)—plays a central role in shaping the structure of matter and driving the dynamics of high-energy collisions. Yet, despite its fundamental importance, QCD poses formidable challenges for theoretical calculations. These challenges become especially pronounced in the analysis of data from high-energy experiments such as the LHC [9] and RHIC [10], as well as from proposed future colliders like the FCC [11], where increasingly precise predictions for QCD processes are essential to disentangle potential signals of new physics from the overwhelming background of strong interactions.

To meet the growing demand for precision, a variety of computational tools—such as Herwig [12–14], Pythia [15], and Sherpa [16]—have been developed to simulate QCD radiation using parton shower algorithms. Remarkable progress has been made in improving the underlying evolution models, with recent milestones including the achievement of next-to-next-to-leading logarithmic (NNLL) accuracy by the PanScales collaboration [17]. However, conventional parton showers remain fundamentally constrained by their reliance on probabilistic frameworks, which inherently neglect quantum interference effects. This thesis is devoted to the development and application of an alternative approach: the amplitude-level parton shower formalism and its Monte Carlo implementation, `CVolver` (Colour eVolver). Building on several years of theoretical work [18–22] and algorithmic development [23, 24], we present the most advanced implementation of `CVolver` to date. Within this framework, we carry out a detailed investigation of subleading colour effects

in perturbative QCD, providing new insight into quantum interference phenomena that lie beyond the reach of traditional parton showers.

This work is structured as follows. Chapter 1 introduces the fundamental theoretical concepts underlying QCD, along with an overview of soft radiation and logarithmic resummation. Chapter 2 defines the general soft evolution algorithm that forms the backbone of the amplitude-level parton shower, `CVolver`. We also introduce the colour flow basis and use it to describe evolution in colour space. In Chapter 3, we present the Monte Carlo implementation of `CVolver`, including its ability to track colour suppression throughout the evolution. Chapter 4 contains a systematic study of subleading colour effects in processes involving multiple soft gluon emissions, and in Chapter 5, we use `CVolver` as an event generator to analyse soft radiation differentially.

Chapter 1

QCD preliminaries

1.1 The QCD Lagrangian

Quantum Chromodynamics (QCD) is the quantum field theory that describes the strong interaction [25, 26] and its associated particles: quarks and gluons [27, 28]. It is a Yang-Mills theory [29], governed by the non-Abelian gauge symmetry group SU(N_c), where N_c is the number of colour charges [30–32]. What follows is a summary of the main building blocks of QCD, and for detailed introductions we refer the reader to [33–35].

The QCD Lagrangian can be decomposed into three distinct parts: the classical Lagrangian, the gauge-fixing term, and a term for the ghost fields:

$$\mathcal{L}_\mathrm{QCD} = \mathcal{L}_\mathrm{classical} + \mathcal{L}_\mathrm{gauge\text{-}fixing} + \mathcal{L}_\mathrm{ghosts}. \tag{1.1}$$

The classical part of the Lagrangian can be separated into the kinetic term for the gluon field, and a term containing the quark fields, their masses, and their interactions with gluons. It is given by

$$\mathcal{L}_\mathrm{classical} = -\frac{1}{4} F^a_{\mu\nu} F_a^{\mu\nu} + \sum_\mathrm{flavours} \bar{q}_i \left(i \not{D} - m \right)_{ij} q_j, \tag{1.2}$$

where $F^a_{\mu\nu}$ is the gluon field strength tensor, and $\not{D}$ is the covariant derivative, where the slashed notation implies contraction with the Dirac matrices [36, 37], $\not{D} = \gamma_\mu D^\mu$. The quark spinor fields describe spin-$\frac{1}{2}$ fermions of mass m, and spinor indices have been suppressed. There are six different quark flavours, with different charges and masses each, that are being summed over. The indices i and j run over the fundamental representation of SU(N_c). In QCD, $N_\mathrm{c} = 3$ and therefore quarks fields are represented by triplet states, sometimes referred to as red, green, and blue. The Greek indices indicate space-time dimensions, with the flat metric $\eta^{\mu\nu} = \mathrm{diag}(1, -1, -1, -1)$. Gluons are in the adjoint representation of SU(N_c), with indices $a, b, c, ...$ that run over $N_\mathrm{c}^2 - 1$ degrees of freedom. The gluon field strength tensor is defined as

$$F^a_{\mu\nu} = \partial_\mu \mathcal{A}^a_\nu - \partial_\nu \mathcal{A}^a_\mu - g_s f^{abc} \mathcal{A}^b_\mu \mathcal{A}^c_\nu, \tag{1.3}$$

where g_s is the coupling constant, f^{abc} are the structure constants of the Lie algebra of $SU(3)$, and $\mathcal{A}^a_\mu$ is the gluon field. In an Abelian gauge theory, $f^{abc} = 0$, forbidding the self-interaction of force-carrying bosons. However, since the SU(N_c) structure constants

are non-zero, the gluon field kinetic term $F^a_{\mu\nu}F^{\mu\nu}_a$ contains 3-point and 4-point gluon vertices. The massless, self-interacting bosons of QCD are behind many of the fundamental properties of QCD, such as colour confinement and asymptotic freedom [38–40].

The covariant derivate takes a different form depending on whether it acts on triplet (quark) or octet (gluon) fields:

$$(D_\mu)_{ij} = \partial_\mu \delta_{ij} + ig_s(t^a \mathcal{A}^a_\mu)_{ij}, \quad (D_\mu)_{ab} = \partial_\mu \delta_{ab} + ig_s(T^c \mathcal{A}^c_\mu)_{ab}, \tag{1.4}$$

where t^a are the generators of SU(N_c) in the fundamental representation, and T^a in the adjoint representation. They have dimensions $N_\mathrm{c} \times N_\mathrm{c}$ and $(N_\mathrm{c}^2-1)\times(N_\mathrm{c}^2-1)$ respectively and satisfy the commutation relations of the Lie algebra [41]:

$$[t^a, t^b] = if^{abc}t^c, \qquad [T^a, T^b] = if^{abc}T^c. \tag{1.5}$$

The generators in the adjoint representation are defined as the structure constants $(T^a)^{bc} = -if^{abc}$. In the fundamental representation they are conventionally represented by the Gell-Mann matrices $t^a = \frac{1}{2}\lambda^a$, which are given by[42]

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$\lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \quad \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \tag{1.6}$$

$$\lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix},$$

which are Hermitian and traceless. The Gell-Mann matrices are normalised so that

$$\mathrm{Tr}\, t^a t^b = T_R\, \delta^{ab} = \frac{\delta^{ab}}{2}. \tag{1.7}$$

Another important element of the Lie algebra are the Casimir operators, which commute with all other generators and characterise each representation. In the fundamental and adjoint representations they are respectively defined:

$$\begin{aligned} \sum_a t^a_{ik} t^a_{kj} &= C_F \delta_{ij}, \qquad & C_F &= \frac{N_\mathrm{c}^2 - 1}{2N_\mathrm{c}} = \frac{4}{3}, \\ \sum_a T^a_{bd} T^a_{dc} &= C_A \delta_{bc}, \qquad & C_A &= N_\mathrm{c} = 3 \end{aligned} \tag{1.8}$$

where C_F and C_A are the Casimir invariants of each representation, which appear frequently in QCD calculations.

As stated at the beginning of this section, QCD is a gauge theory of SU(N_c), and therefore is invariant under the local gauge transformations of this Lie group. Local means the redefinition of the fields can vary across each point in spacetime without altering the physical quantities. In the case of the quark fields the gauge transformation is

$$q_i(x) \to q'_i(x) = \exp(it^a\theta^a(x))_{ij} q_j(x) \equiv \Omega(x)_{ij} q_j(x), \tag{1.9}$$

where $\Omega(x) \equiv \exp(it^a\theta^a(x))$ is the gauge transformation as a function of the 4-dimensional spacetime coordinate x. The covariant derivate also transforms like the quark fields:

$$D^\mu_{ij} q_j(x) \to D'^\mu_{ij} q'_j(x) \equiv \Omega_{ij}(x) D^\mu_{jk} q_k(x), \tag{1.10}$$

which requires that the gluon fields transform like:

$$t^a \mathcal{A}^a_\mu \to t^a \mathcal{A}'^a_\mu = \Omega(x) t^a \mathcal{A}^a_\mu \Omega^{-1}(x) + \frac{i}{g_s} (\partial_\mu \Omega(x)) \Omega^{-1}(x). \tag{1.11}$$

From this result, the transformation of the gluon field strength tensor can be derived:

$$t^a F^a_{\mu\nu} \to t^a F'^a_{\mu\nu} = \Omega(x) t^a F^a_{\mu\nu} \Omega^{-1}(x), \tag{1.12}$$

and therefore, in contrast to the QED electromagnetic field tensor, it is not gauge invariant. This is another consequence of the non-Abelian quality of SU(N_c). Combining the trans-

formations given in Eqs. (1.9), (1.10), and (1.12), it can be shown that the classical part of the Lagrangian in Eq. (1.2) is gauge invariant, which is necessary for the renormalizability of the theory [43–46]. QCD is also invariant under many global symmetries, some of them exact, like the Poincaré or U(1) groups, and some of them approximate, like the light quark flavour symmetries. For a more thorough discussion of the symmetries within QCD, see [34].

So far, we have considered the QCD Lagrangian within the framework of classical field theory. However, to perform perturbative calculations [47], we must transition to a quantum field theory (QFT) description. In QFT, fields are elevated to operators that act on states in Fock space through a process called canonical quantisation [48]. This is analogous to how position and momentum become operators in quantum mechanics. Consequently, the Lagrangian itself becomes an operator. It separates into two components: a free part, which governs the independent propagation of particles, and an interaction term, which we handle using perturbation theory [33, 49].

In a quantised theory, Feynman rules emerge as a systematic way to compute scattering amplitudes. We will provide an abridged overview of how this is done, but for details see [33, 35]. The key quantity is the S-matrix, which relates the initial and final asymptotic states of a scattering process. Formally, the S-matrix is constructed from the time evolution operator in the interaction picture, expressed as a time-ordered exponential of the interaction Lagrangian. Expanding this exponential in a perturbative series and inserting a complete basis of states leads to a sum of integrals over intermediate virtual particles. These integrals are naturally represented as Feynman diagrams, where each term in the expansion corresponds to a distinct interaction process. The Feynman rules, which provide the mathematical expressions for these diagrams, follow directly from the structure of the quantised Lagrangian: propagators arise from the free part, while vertices and interaction terms originate from the perturbative expansion of the interacting part.

Attempting to derive a propagator for the gluon from the Lagrangian in Eq. (1.2) is impossible due to gauge invariance. It is necessary to add a gauge fixing term:

$$\mathcal{L}_{\text{gauge-fixing}} = -\frac{1}{2\xi}(\partial^\mu \mathcal{A}^a_\mu)^2, \tag{1.13}$$

which defines the set of covariant gauges, so called because they are Lorentz invariant. The gauge-fixing parameter ξ can take different values: $\xi = 1$ corresponds to the Feynman gauge, and $\xi = 0$ corresponds to the Lorenz or Landau gauge. Due to the overall gauge invariance of the theory, physical observables, specifically the S-matrix, will be

independent of ξ. With this set of gauges, the gluon propagator can be derived:

$$\Delta^{\mu\nu,ab}(p) = \frac{i\delta^{ab}}{p^2 + i\epsilon}\left(-\eta^{\mu\nu} + (1-\xi)\frac{p^\mu p^\nu}{p^2 + i\epsilon}\right), \tag{1.14}$$

where the second term is null in the Feynman gauge. However, the price to pay for using a covariant gauge is that a set of unphysical, ghost particles must be included, to restore global BRST invariance [50, 51]. They are included with an additional term to the Lagrangian:

$$\mathcal{L}_{\text{ghosts}} = (\partial_\mu \bar{c}^a) D^\mu_{ab} c^b, \tag{1.15}$$

where $\bar{c}^a$ and c^a are the the Faddeev-Popov ghost fields, and they are anti-commuting scalars [44, 52]. Without these unphysical degrees of freedom, QCD with a covariant gauge loses gauge invariance and cannot be renormalised. There is an alternative gauge-fixing term which results in the ghosts decoupling from physical processes:

$$\mathcal{L}_{\text{gauge-fixing}} = -\frac{1}{2\xi}\left(n^\mu \mathcal{A}^a_\mu\right)\left(n^\nu A^a_\nu\right), \tag{1.16}$$

and these are called the axial gauges, where n^μ fixes some direction in spacetime. The gluon propagator with this gauge is given by

$$\Delta^{\mu\nu,ab}(p) = \frac{i\delta^{ab}}{p^2 + i\epsilon}\left(-\eta^{\mu\nu} + \frac{p^\mu n^\nu + n^\mu p^\nu}{p\cdot n} - (n^2 + \xi p^2)\frac{p^\mu p^\nu}{(p\cdot n)^2}\right), \tag{1.17}$$

which has more terms than Eq. (1.14) and divergences at $p \cdot n = 0$. The case where $\xi = 0$ and n^μ is light-like is called the light-cone gauge [53], and it causes the last term to vanish. These gauges can still lead to complicated calculations, especially if the $p \cdot n$ pole needs to cancel in numerical simulations.

We end here our overview of the QCD Lagrangian and its properties as a quantised gauge theory. A complete set of Feynman rules can be found in many sources, for example [33–35]. We continue by considering a simple perturbative calculation.

1.2 The $q'\bar{q} \to q'\bar{q}$ Born amplitude

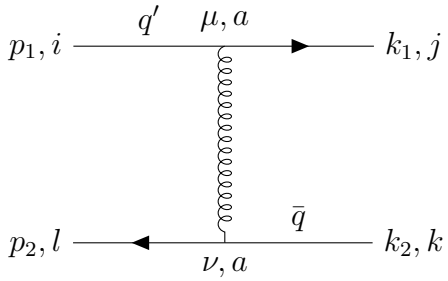


Figure 1.1: Feynman diagram contributing to the $q'\bar{q} \to q'\bar{q}$ process in tree-level QCD. The primes indicate different flavours. The four external fermions are labelled by their momenta and fundamental representation index. The quark-gluon vertices are labelled by the spacetime and adjoint representation indices of the gluon.

We begin by considering the Born amplitude of the $q'\bar{q} \to q'\bar{q}$ process. Due to the different flavours, the only contributing diagram is t-channel gluon exchange, shown in Fig. 1.1. We perform the calculation in the high energy limit, so all the fermions are massless. The amplitude is given by

$$\mathcal{M} = \bar{u}_j(k_1)(-ig_s t^a_{ji}\gamma^\mu)u_i(p_1)\left(\frac{-i\eta_{\mu\nu}}{q^2 + i\epsilon}\right)v_k(k_2)(-ig_s t^a_{kl}\gamma^\nu)\bar{v}_l(p_2), \tag{1.18}$$

where $u_i(p_1)$ is the Dirac spinor of the incoming quark, $\bar{u}_j(k_1)$ of the outgoing quark, $v_k(k_2)$ of the outgoing anti-quark, and $\bar{v}_l(p_2)$ of the incoming anti-quark. The exchanged gluon has momentum $q = (p_1 - k_1)$. Spinor indices have been suppressed. The gluon propagator uses the covariant Feynman gauge $\xi = 1$, defined in Eq. (1.14). We do not need to worry about Faddeev-Popov ghosts because they do not contribute at tree level [33, 44]. We can square the amplitude, sum over all spins and colours, and average over the initial spins and colours to obtain:

$$\begin{aligned}\frac{1}{9}\sum_{\text{colours}}\frac{1}{4}\sum_{\text{spins}}|\mathcal{M}|^2 &= \frac{1}{9}\,\text{Tr}[t^b t^a]\,\text{Tr}[t^b t^a] \times \frac{g_s^4}{4q^4}\,\text{Tr}[\not{k}_1\gamma^\mu\not{p}_1\gamma^\nu]\,\text{Tr}[\not{k}_2\gamma_\mu\not{p}_2\gamma_\nu], \\ &= \frac{16g_s^4}{9q^4}[(k_1\cdot k_2)(p_1\cdot p_2) + (k_1\cdot p_2)(p_1\cdot k_2)]\end{aligned} \tag{1.19}$$

and we note the colour part has factorised from the spin and kinematics. The second line uses the relation $\text{Tr}[t^a t^b] = \delta^{ab}/2$ to evaluate the colour traces. We have also used the

completeness relations for the Dirac spinors [33, 35]:

$$\sum_s u^s(p)\bar{u}^s(p) = \not{p} + m \qquad \sum_s v^s(p)\bar{v}^s(p) = \not{p} - m, \tag{1.20}$$

to obtain the spinor traces. These can be evaluated using standard identities [33, 35]. It is useful to rewrite scattering amplitudes in terms of the Mandelstam variables [54], which are defined:

$$\begin{aligned} s &= (p_1 + p_2)^2 = (k_1 + k_2)^2 = 2(p_1 \cdot p_2) \\ t &= (p_1 - k_1)^2 = (k_2 - p_2)^2 = -2(p_1 \cdot k_1) \\ u &= (p_1 - k_2)^2 = (k_1 - p_2)^2 = -2(p_1 \cdot k_2) \end{aligned} \tag{1.21}$$

where the final equality works in the massless limit of the fermions. In the centre of mass frame of the system, $s = (p_1 + p_2)^2 = E_{\text{cm}}^2$, where E_{cm} is the total energy of the system in this frame. The scattering amplitude in Eq. (1.19) then reads

$$\frac{1}{9}\sum_{\text{colours}}\frac{1}{4}\sum_{\text{spins}}|\mathcal{M}|^2 = \frac{16g_s^4}{9t^2}\left[\left(\frac{s}{2}\right)^2 + \left(\frac{u}{2}\right)^2\right] \tag{1.22}$$

which shows this contribution scales as $s/t = (p_1 \cdot p_2)/q^2$ when $s \gg -t$. Due to crossing symmetry [33, 55], any $2 \to 2$ diagram with a virtual particle exchange will have a corresponding scaling behaviour, obtained by exchanging the Mandelstam variables. For example, the s-channel gluon exchange scattering contributing to $q'\bar{q}' \to q\bar{q}$ contains a term $(t^2 + u^2)/s^2$.

We shall go back to the Born scattering amplitude in Eq. (1.19), and consider an alternative way of evaluating the colour factor. This might seem a little long-winded, but it will provide a connection to the colour structures used for the rest of this thesis. We will employ the Fierz identity:

$$\sum_a t^a_{ij}t^a_{kl} = \frac{1}{2}\left(\delta_{il}\delta_{kj} - \frac{1}{N_{\text{c}}}\delta_{ij}\delta_{kl}\right), \tag{1.23}$$

which separates the gluon propagator connected to two fermion lines in Fig. 1.1 into two terms contracting the four fundamental colour indices pairwise. These two terms can be

understood as two different ways for the colour to "flow" in the diagram. If we apply the Fierz identity twice to the squared amplitude, we obtain

$$\begin{aligned}\sum_a t^a_{ij}t^a_{kl}\sum_b t^b_{ji}t^b_{lk} =& \frac{1}{4}\Big[(\delta_{il}\delta_{jk})(\delta_{il}\delta_{jk}) - \frac{1}{N_\mathrm{c}}(\delta_{il}\delta_{jk})(\delta_{ij}\delta_{lk}) \\ & - \frac{1}{N_\mathrm{c}}(\delta_{ij}\delta_{lk})(\delta_{il}\delta_{jk}) + \frac{1}{N_\mathrm{c}^2}(\delta_{ij}\delta_{lk})(\delta_{ij}\delta_{lk})\Big],\end{aligned} \tag{1.24}$$

where in each of the four terms the first parenthesised quantity contains the flow of fundamental colour indices in the amplitude, and the second parenthesised quantity contains the flow in the conjugate amplitude. Thus we can write the squared amplitude as a density matrix:

$$M = \frac{H}{4}\begin{pmatrix} 1 & -1/N_c \\ -1/N_c & 1/N_c^2 \end{pmatrix}, \quad H = \frac{8g_s^4}{9t^2}\left[\left(\frac{s}{2}\right)^2 + \left(\frac{u}{2}\right)^2\right], \tag{1.25}$$

where H contains the kinematic and spin contributions to the squared amplitude, and each of the four elements in the matrix corresponds to the possible ways to connect the colour indices in the amplitude and conjugate amplitude, as per Eq. (1.24). We are working in a vector space of dimension 2, each corresponding to the two terms of the Fierz identity in Eq. (1.23). We also need to evaluate the Kronecker delta contractions that appear in Eq. (1.24):

$$S = \begin{pmatrix} (\delta_{il}\delta_{jk})(\delta_{il}\delta_{jk}) & (\delta_{il}\delta_{jk})(\delta_{ij}\delta_{lk}) \\ (\delta_{il}\delta_{jk})(\delta_{ij}\delta_{lk}) & (\delta_{ij}\delta_{lk})(\delta_{ij}\delta_{lk}) \end{pmatrix} = \begin{pmatrix} N_\mathrm{c}^2 & N_\mathrm{c} \\ N_\mathrm{c} & N_\mathrm{c}^2 \end{pmatrix}, \tag{1.26}$$

where we have defined S, the scalar product matrix that contains the contractions between amplitude and conjugate amplitude for each combination of colour flows. Thus, the full result of the scattering process, which is evaluated by performing the sum in Eq. (1.24), can be written in this notation as

$$\mathrm{Tr}[MS] = \sum_{\sigma,\tau} M_{\sigma\tau}S_{\tau\sigma} = \frac{N_\mathrm{c}^2 - 1}{4}H = \frac{16g_s^4}{9t^2}\left[\left(\frac{s}{2}\right)^2 + \left(\frac{u}{2}\right)^2\right], \tag{1.27}$$

and we have arrived to the same result as in Eq. (1.22). The labels σ, τ run over each colour flow configuration, and in this case they take values 1 or 2, corresponding to each

term in the Fierz identity. This is our first encounter with the colour flow basis [20, 56, 57], which will be discussed in depth in Chapter 2, and will remain important for the rest of this work. Now we return to study QCD radiation in the soft limit.

1.3 The eikonal approximation

Experiments at hadron colliders produce a large amount of QCD radiation, which is dominated by the points in phase-space where amplitudes become large or diverge. Therefore, to make meaningful phenomenological predictions, we often need to account for soft and collinear radiation beyond the Born amplitude. In this work we focus on the soft limit, where we rescale the emitted momenta:

$$k^\mu \to \lambda k^\mu \tag{1.28}$$

where λ is a scalar quantity. In the limit $\lambda \to 0$ all components of the emitted momentum are vanishingly small relative to the hard partons. Taking this limit and keeping only the leading terms in λ is the eikonal approximation [58, 59].

We will consider only the emission of soft gluons, because the $g \to q\bar{q}$ splittings are suppressed in the eikonal limit. We start with a general amplitude, $\mathcal{M}_n(\{p_n\})$, which contains n external hard partons with momenta $\{p_n\}$. Introducing an additional soft gluon with momentum λk, emitted from an outgoing quark with momentum p, modifies the amplitude as

$$\begin{aligned}
\mathcal{M}^q_{n+1} &= \bar{u}_i(p)\left[\left(-ig_s t^a_{ij}\gamma^\mu \epsilon^{*a}_\mu(\lambda k)\right)\frac{i(\not{p}+\lambda\not{k})}{(p+\lambda k)^2+i\epsilon}\right] \\
&\quad\times \mathcal{M}_{n,j}(p_1,\ldots,p+\lambda k,\ldots,p_n) \\
&= \bar{u}_i(p)\left[\left(-ig_s t^a_{ij}\epsilon^{*a}_\mu(\lambda k)\right)\frac{i(2p^\mu+2\lambda k^\mu-\lambda\not{k}\gamma^\mu)}{(p+\lambda k)^2+i\epsilon}\right] \\
&\quad\times \mathcal{M}_{n,j}(p_1,\ldots,p+\lambda k,\ldots,p_n)
\end{aligned} \tag{1.29}$$

where $\mathcal{M}^q_{n+1}$ is the amplitude for a soft gluon connected to an external quark. We have extracted from $\mathcal{M}_{n,j}$ the spinor associated with the external quark from which the emission occurs, $\bar{u}_i(p)$, and we are working in the massless limit of the quarks. After taking the soft

limit we obtain

$$\lim_{\lambda\to 0}\mathcal{M}^{f}_{n+1}=\frac{\bar{u}_i(p)}{\lambda}\epsilon^{*a}_{\mu}(k)\left[\frac{g_s t^a_{ij}p^{\mu}}{p\cdot k+i\epsilon}\right]\mathcal{M}_{n,j}(p_1,\ldots,p,\ldots,p_n)+\mathcal{O}(\lambda^0), \tag{1.30}$$

which shows that the emission of the soft gluon factorises from the rest of the amplitude, and it decouples from the spin of the emitter. We identify the factor $(g_s t^a_{ij}p^{\mu})/(p\cdot k+i\epsilon)$ as the effective Feynman rule of the emission. In the case of a soft gluon emitted from an anti-quark, the Feynman rule is equivalent, but with an opposite sign. Emissions from off-shell internal lines are suppressed, so we only consider external legs. This follows from the denominator of the associated propagator, $(p+\lambda k)^2 \xrightarrow[\lambda\to 0]{} p^2 \neq 0$ if the parton with momentum p is off-shell [34].

An equivalent rule can be obtained for the emission of a soft gluon from an external gluon line. The derivation, using the light-cone gauge, can be found in [34, 60], and the final result is

$$\lim_{\lambda\to 0}\mathcal{M}^{g}_{n+1}=\frac{\epsilon^{*b}_{\mu}(p)}{\lambda}\epsilon^{*a}_{\nu}(k)\left[\frac{g_s i f^{abc}p^{\nu}}{p\cdot k+i\epsilon}\right]\mathcal{M}^{c,\mu}_{n}(\{p_n\})+\mathcal{O}(\lambda^0) \tag{1.31}$$

where the emission factorises from the amplitude again and we identify another effective Feynman rule, $(g_s i f^{abc}p^{\nu})/(p\cdot k+i\epsilon)$. Note that the only difference between the Feynman rules in Eqs. (1.30) and (1.31) is the colour structure at the vertex. Armed with the effective Feynman rules in the eikonal approximation, we can evaluate the contribution of one soft emission to any amplitude.

1.3.1 One real emission

The leading contribution in the soft limit is found by attaching a soft gluon to the external legs. Therefore, to consider all possible contributions to a real soft emission we need to sum over the external legs i:

$$\mathbf{D}|\mathcal{M}_n\rangle=-\sum_{\lambda}\sum_{i}g_s\mathbf{T}^a_i\frac{\epsilon^{*a}_{\lambda}\cdot p_i}{p_i\cdot k+i\epsilon}|\mathcal{M}_n\rangle, \tag{1.32}$$

where λ is the polarisation of the emitted gluon and $|\mathcal{M}_n\rangle$ is the tree-level amplitude for the scattering of n on-shell massless quarks and gluons. We have defined the colour charge

operator $\mathbf{T}_i^a$, which in the Gell-Mann basis will take a different form depending on the parton i, as shown in Eqs. (1.30) and (1.31):

$$(\mathbf{T}_i^a)_{dc} = \begin{cases} t_{dc}^a & i = \bar{u} \\ t_{dc}^a & i = \bar{v} \\ -t_{cd}^a & i = u \\ -t_{dc}^a & i = v \\ if^{dac} & i = \epsilon \\ -if^{cad} & i = \epsilon^* \end{cases} \tag{1.33}$$

and we have omitted the colour indices in $|\mathcal{M}_n\rangle$ that contract with $\mathbf{T}_i^a$.

The amplitude for soft gluon emission exhibits divergences in two distinct regions. The first occurs when the energy of the emission becomes vanishingly small, $k \to 0$. This is the soft divergence. The second occurs when the emitted gluon becomes collinear with any of the external legs, $p_i \cdot k \to 0$. These are collinear divergences. These are the two types of infrared (IR) divergences, which occur because the amplitude $|\mathcal{M}_{n+1}\rangle$ becomes indistinguishable from $|\mathcal{M}_n\rangle$ when the emission has zero energy or is completely collinear with another leg [61, 62]. We can cure these divergences by including all diagrams that contribute at the same order in α_s, and so we proceed by considering a soft virtual loop.

1.3.2 One virtual loop

We want to evaluate the operator that dresses the tree-level amplitude $|\mathcal{M}_n\rangle$ with all possible virtual soft gluon exchanges. A rigorous treatment is quite involved. The calculations can be found in [58, 61–65], and they show that self-energy corrections, quark loops and loops involving ghosts are all subleading in the soft limit. Contour integrals must also be performed for the three leading topologies: virtual gluon exchange between two fermion lines, a fermion and a gluon, or two gluons, arriving at the final result:

$$\mathbf{\Gamma}|\mathcal{M}_n\rangle = \frac{\alpha_s}{\pi} \sum_{i<j} (-\mathbf{T}_i^a \cdot \mathbf{T}_j^a) \int \frac{\mathrm{d}E_k}{E_k} \left(\int \frac{\mathrm{d}\Omega_k}{4\pi} \omega_{ij}(n_k) - i\pi \bar{\delta}_{ij} \right) |\mathcal{M}_n\rangle, \tag{1.34}$$

where i and j are distinct external legs in $|\mathcal{M}_n\rangle$, and $\bar{\delta}_{ij}$ is 1 if i and j are both incoming or both outgoing, and zero otherwise. The imaginary term $i\pi\bar{\delta}_{ij}$ is the Coulomb contribution

and we will return to it next section. The antenna function is defined:

$$\omega_{ij}(n_k) = E_k^2 \frac{p_i \cdot p_j}{(p_i \cdot k)(p_i \cdot k)}, \tag{1.35}$$

which diverges if the energy of the emitted gluon vanishes, or if it becomes collinear with i or j. This is the same infrared divergence structure that plagued the soft emission in Eq. (1.32). In fact, when computing the fully inclusive correction to the amplitude from a soft gluon, including both real and virtual contributions, there is total cancellation:

$$\mathrm{Re}[\langle \mathcal{M}_n | \mathbf{\Gamma} | \mathcal{M}_n \rangle] + \mathrm{Re}[\langle \mathcal{M}_n | \mathbf{\Gamma}^\dagger | \mathcal{M}_n \rangle] + \int \mathrm{d}\Pi_{n+1} \langle \mathcal{M}_n | \mathbf{D}^\dagger \mathbf{D} | \mathcal{M}_n \rangle = 0, \tag{1.36}$$

where $\mathrm{d}\Pi_{n+1}$ denotes the phase-space element of the real emission. This cancellation is not unexpected; it follows from the Kinoshita-Lee-Nauenberg (KLN) theorem, which states that all IR divergences cancel perturbatively at each order in α_s [66, 67]. An equivalent result was proven by Bloch and Nordsieck for QED [68]. To cancel every IR pole perturbatively, all degenerate states must be included. In this context, degenerate means an amplitude with any number of soft or collinear gluons, which are physically indistinguishable from the tree-level amplitude.

We have not considered UV divergences. In this work we always work in the soft limit, so UV divergences are regulated by the hard scale of the scattering. Nevertheless, if we evaluated virtual corrections to the Born amplitude in Eq. (1.19) without the eikonal approximation, we would obtain UV divergent loop integrals [69]. These poles are removed through renormalisation, and are systematically reabsorbed into a finite number of parameter redefinitions in the Lagrangian [43, 70–72].

1.4 The gaps-between-jets observable

For a fully-inclusive observable, soft virtual and real corrections cancel perturbatively, as per the KLN theorem. However, this is not the case if a restriction is placed on the phase-space of the real emissions spoiling the cancellation. We still require that the divergent poles cancel, and observables which preserve this quality are named infrared and collinear (IRC) safe [34]. Nevertheless, an IRC observable will generally incur finite and potentially large logarithmic corrections due to soft and collinear radiation if the observable is not fully-inclusive. These large logarithms L can spoil the perturbative expansion if

they compensate for the smallness of the coupling constant α_s, thus requiring resummation at all orders. We reorganise the perturbative series [34]:

$$\frac{\Sigma}{\Sigma_0} = \sum_{n=0}^{\infty} C_0^n \alpha_s^n L^{2n} + \sum_{n=1}^{\infty} C_1^n \alpha_s^n L^{2n-1} + \sum_{n=1}^{\infty} C_2^n \alpha_s^n L^{2n-2} + \dots, \tag{1.37}$$

where Σ is an observable sensitive to soft and collinear enhancements, Σ_0 is the contribution from the Born amplitude, and $C_0^0 = 1$. The first term corresponds to the resummed leading log (LL) contribution, the second term is the next-to-leading log (NLL), and so forth.

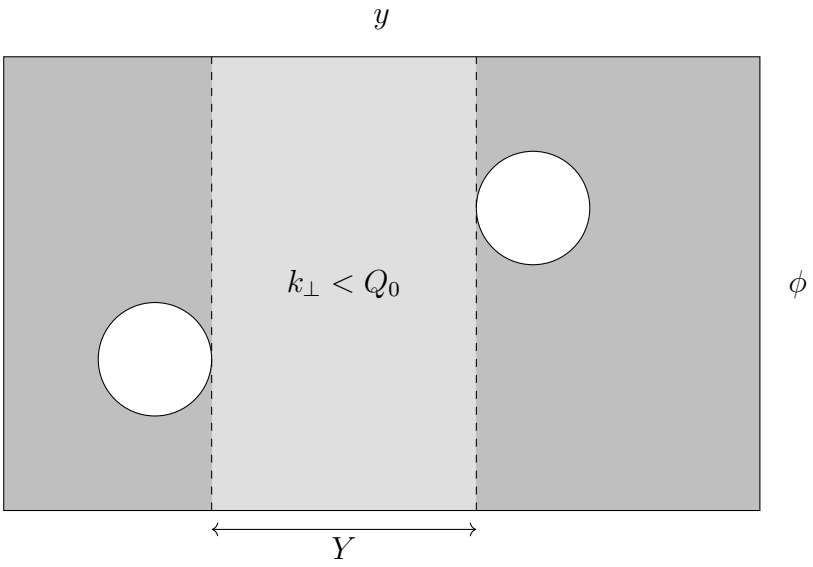


Figure 1.2: The gaps-between-jets observable. The white circles are the two hard jets, and radiation in the interjet rapidity region is constrained to $k_\perp < Q_0$, where Q_0 is the inclusivity scale.

For most of this work we will focus on the gaps-between-jets observable, which vetoes radiation above some scale in a "gap" region. It is depicted in Fig. 1.2, where y is rapidity and ϕ is the azimuth. The two primary jets are represented with two white circles, and radiation in the interjet region is restricted to $k_\perp < Q_0$. Applying such a veto to suppress radiation in the rapidity gap between two jets is of significant phenomenological importance, as it provides sensitivity to the underlying colour structure of the hard scattering process. For example, in the vector boson fusion (VBF) production of the Higgs boson, the initial-state partons exchange electroweak gauge bosons in a colour-singlet configuration. As a consequence, the QCD radiation in the rapidity gap between the two final-state quarks is suppressed. This clean rapidity gap is exploited experimentally to distinguish VBF events from background processes involving colour-octet exchange, such as gluon

fusion or QCD-induced dijet production, where additional radiation is more likely to populate the gap. The absence of central jet activity is routinely used as a tagging criterion in experimental analyses of VBF Higgs production, for example in [73].

The gaps-between-jets observable is sensitive to wide-angle soft gluon emissions and inclusive over collinear radiation. Therefore, it is single logarithmic:

$$\frac{\Sigma}{\Sigma_0} = \sum_{n=0}^{\infty} C_0^n \alpha_s^n L^n + \sum_{n=1}^{\infty} C_1^n \alpha_s^n L^{n-1} + \sum_{n=1}^{\infty} C_2^n \alpha_s^n L^{n-2} + \dots . \tag{1.38}$$

The summation of these effects has a long history and is of considerable theoretical interest [74–88]. The original calculation was performed by Oderda and Sterman in 1998 [75, 76]. They used the eikonal approximation to dress the $qq \to qq$, t-channel scattering with all possible soft virtual gluons, so as to sum the leading logarithms:

$$|\mathcal{M}\rangle = \mathbf{V}_{Q_0,Q}|\mathcal{M}_0\rangle, \quad \mathbf{V}_{a,b} = \exp\left[\frac{\alpha_s}{\pi}\int_a^b \frac{\mathrm{d}k_\perp}{k_\perp}\mathbf{\Gamma}\right], \tag{1.39}$$

where $|\mathcal{M}_0\rangle$ is the tree-level scattering, Q is the hard scale of the process, and Q_0 is the veto scale. The lower limit of the evolution can be set to the veto scale due to KLN cancellation, since the observable is inclusive over all radiation below that scale. The soft anomalous dimension is

$$\mathbf{\Gamma} = \int_{\text{gap}} \mathrm{d}y \frac{\mathrm{d}\phi}{2\pi} \sum_{i<j} \mathbf{T}_i \cdot \mathbf{T}_j \, \omega_{ij} - 2\pi i \, \mathbf{T}_1 \cdot \mathbf{T}_2, \tag{1.40}$$

which is the k_T ordered form of Eq. (1.34). The angular integration is over the gap region, since KLN tells us again the virtuals will cancel with the reals out of gap. The incoming partons are labelled 1 and 2, while the outgoing are 3 and 4. The logarithmic contribution to the observable is found by squaring the evolved amplitudes:

$$\Sigma_0 = \langle \mathcal{M}|\mathcal{M}\rangle = \langle \mathcal{M}_0|\mathbf{V}^\dagger_{Q_0,Q}\mathbf{V}_{Q_0,Q}|\mathcal{M}_0\rangle = \mathrm{Tr}(\mathbf{V}_{Q_0,Q} H \mathbf{V}^\dagger_{Q_0,Q}). \tag{1.41}$$

It may seem this is enough to capture correctly all leading logarithmic contributions, but in 2001 Dasgupta and Salam spotted the error [86]. Figure 1.3 depicts the problem. In

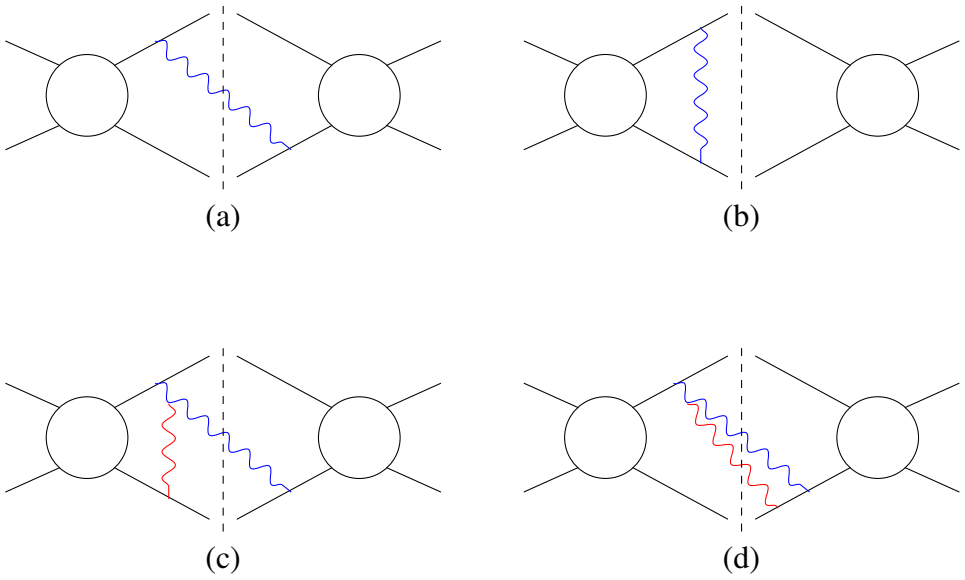


Figure 1.3: Depiction of the miscancellations leading to non-global logarithms. The blue lines represent gluons out of the gap, while the red lines represent those in the gap. All are above the inclusivity scale Q_0.

(a) there is a single gluon being emitted into the out-of-gap region, which cancels with the out-of-gap virtual in (b). Therefore, for a single emission it is safe to discard all emissions out of gap. Now we consider 2 gluons, in (c) and (d). In (c) the first gluon is emitted out of gap, and then a virtual gluon is exchanged with momentum in the gap region between the emitted gluon and a hard jet. In (d) the first gluon is the same out-of-gap emission, and the second gluon is a real emission into the gap region. The veto forbids (d), but allows (c), so that contribution survives and creates a new source of leading logarithms. Therefore, we must dress the hard scattering with any number of out-of-gap emissions, and dress all of those with any number of virtual exchanges. The contributions from these miscancellations are termed non-global logarithms, as opposed to the global logarithms evaluated in Eq. (1.41). The existence of non-global logarithms greatly increases the challenge of resumming soft gluon corrections.

The story did not end there. In 2006, Forshaw, Kyrieleis, and Seymour found another source of unaccounted-for logarithms, which arise from the imaginary parts in the anomalous dimension [87, 88]. They occur at lowest order α_s^4 relative to the Born process, and they were named the super-leading logarithms because at this order they go as $\sim \alpha_s^4 \log^5 Q/Q_0$, which is super leading compared to the leading single logarithmic contribution $\sim \alpha_s^n \log^n Q/Q_0$. There is a lot of on-going work being done to evaluate these terms and their effects [89–100].

Chapter 2

Amplitude evolution

The following chapter covers the general soft gluon evolution algorithm described in [20], as well as results from [18, 22] which are necessary to understand the evolution in colour space. The algorithm constitutes a special case of the more general approach given in [19].

2.1 The general algorithm

We start by outlining the general evolution algorithm which computes the differential contribution of multiple soft-gluon emissions to a given observable. The precision achieved by the algorithm is dependent on the observable, which is discussed in [19]. This is the structure on which the `CVolver` Monte Carlo is built. The differential cross-section after n gluon emissions is

$$\begin{aligned}
\sigma_0 &= \mathrm{Tr}\left(\mathbf{V}_{\mu,Q}\mathbf{H}(Q)\mathbf{V}^\dagger_{\mu,Q}\right) \equiv \mathrm{Tr}\,\mathbf{A}_0(\mu) \\
\mathrm{d}\sigma_1 &= \mathrm{Tr}\left(\mathbf{V}_{\mu,E_1}\mathbf{D}^\mu_1\mathbf{V}_{E_1,Q}\mathbf{H}(Q)\mathbf{V}^\dagger_{E_1,Q}\mathbf{D}^\dagger_{1\mu}\mathbf{V}^\dagger_{\mu,E_1}\right)\mathrm{d}\Pi_1 \\
&\equiv \mathrm{Tr}\,\mathbf{A}_1(\mu)\mathrm{d}\Pi_1, \\
\mathrm{d}\sigma_2 &= \mathrm{Tr}\left(\mathbf{V}_{\mu,E_2}\mathbf{D}^\nu_2\mathbf{V}_{E_2,E_1}\mathbf{D}^\mu_1\mathbf{V}_{E_1,Q}\mathbf{H}(Q)\mathbf{V}^\dagger_{E_1,Q}\mathbf{D}^\dagger_{1\mu}\mathbf{V}^\dagger_{E_2,E_1}\mathbf{D}^\dagger_{2\nu}\mathbf{V}^\dagger_{\mu,E_2}\right)\mathrm{d}\Pi_1\mathrm{d}\Pi_2 \\
&\equiv \mathrm{Tr}\,\mathbf{A}_2(\mu)\mathrm{d}\Pi_1\mathrm{d}\Pi_2 \\
&\vdots \\
\mathrm{d}\sigma_n &= \mathrm{Tr}\,\mathbf{A}_n(\mu)\prod_{i=1}^{n}\mathrm{d}\Pi_i
\end{aligned} \tag{2.1}$$

where $\mathbf{H}(Q)$ is the hard scattering matrix at a hard scale Q, given by $\mathbf{H} = |\mathcal{M}\rangle\langle\mathcal{M}|$. The vector $|\mathcal{M}\rangle$ is the scattering amplitude. For the simple processes we consider it is just a vector in colour space. Eq. (2.1) is a Markovian process which dresses the hard scattering matrix with virtual evolution operators $\mathbf{V}$ and real emission operators $\mathbf{D}$, the latter increasing the dimension in colour space. The final virtual operator evolves down to an infrared cutoff μ. It is possible to reformulate this algorithm to make the cancellation of infrared divergences explicit at each step, as done in [20].

We continue by defining each of the operators in Eq. (2.1). The real emission operator,

in the eikonal limit, is given by

$$\mathbf{D}_i^\mu = \sum_j \mathbf{T}_j E_i \frac{p_j^\mu}{p_j \cdot q_i} = \sum_j \mathbf{T}_j \frac{n_j^\mu}{n_j \cdot n_i} \tag{2.2}$$

for the emission of a gluon i with energy E_i and momentum q_i. The sum over partons j is context specific: it includes the partons from the hard process as well as all prior gluon emissions. Similarly, the colour charge operators $\mathbf{T}_j$ depend on the representation in $\mathrm{SU}(3)_c$ at each step. These are rectangular operators in colour; by adding another gluon, they increase the size of the colour space. We will discuss what representation we choose in Section 2.1.1. The second equality uses normalized four vectors, $n_i^\mu = p_i^\mu / E_i$, making explicit the fact that the soft limit of $\mathbf{D}_i^\mu$ does not depend on the scale of the emission.

The phase-space element associated with the emission of gluon i is

$$\mathrm{d}\Pi_i = -\frac{\alpha_s}{\pi} \frac{\mathrm{d}E_i}{E_i} \frac{\mathrm{d}\Omega_i}{4\pi} \tag{2.3}$$

where α_s is the QCD coupling constant, and $\mathrm{d}\Omega_i = \sin\theta_i \mathrm{d}\theta_i \mathrm{d}\phi_i$ is the solid angle element, with polar angle θ_i and azimuthal angle ϕ. The frame of reference used in the Monte Carlo implementation is discussed in Section 3.1.1.

The soft-gluon anomalous dimension operator corresponds to dressing all possible pairs of partons at any given step with a single virtual soft gluon exchange, and will serve as the fundamental building block of the Sudakov operator. It is

$$\mathbf{\Gamma} = \frac{\alpha_s}{\pi} \sum_{i<j} (-\mathbf{T}_i \cdot \mathbf{T}_j) \left(\int \frac{\mathrm{d}\Omega_k}{4\pi} \omega_{ij}(n_k) - i\pi \bar{\delta}_{ij} \right) \tag{2.4}$$

where $\bar{\delta}_{ij}$ is defined such that $\bar{\delta}_{ij} = 1$ if partons i and j are both in the initial state or both in the final state, and zero otherwise. The final term corresponds to Coulomb gluon exchange. The sum over parton pairs is also context specific, running over emitted soft gluons as well as partons in the hard process. The colour charge products are also in a context specific representation of $\mathrm{SU}(3)_c$. The angular integral is a function of the normalised four vector n_k, meaning it only depends on its direction.

The dipole, or antenna, function is defined

$$\omega_{ij}(n_k) = E_k^2 \frac{p_i \cdot p_j}{(p_i \cdot k)(p_i \cdot k)} = \frac{n_i \cdot n_j}{(n_i \cdot n_k)(n_j \cdot n_k)}, \tag{2.5}$$

where k is the momentum of the emitted gluon. It is only a function of the direction of the momenta of i, j and the virtual gluon with normalized momentum n_k. Note the divergence as n_k becomes collinear with n_i or n_j. As we discussed in the previous chapter, due to the KLN theorem, these infrared divergences will cancel perturbatively with corresponding collinear divergences in the real emissions, present in Eq. (2.2), when integrated over phase-space.

These soft loops can be resummed, by considering diagrams with n virtual soft gluon exchanges, and summing over all n in all possible orderings of gluon exchange. This results in the Sudakov operator

$$\begin{aligned} \mathbf{V}_{a,b} &= \mathrm{P}\exp\left(-\int_a^b \frac{\mathrm{d}E_k}{E_k}\mathbf{\Gamma}\right) \\ &= \mathrm{P}\exp\left(-\frac{\alpha_s}{\pi}\int_a^b \frac{\mathrm{d}E_k}{E_k}\sum_{i<j}(-\mathbf{T}_i\cdot\mathbf{T}_j)\left[\int\frac{\mathrm{d}\Omega_k}{4\pi}\omega_{ij}(n_k) - i\pi\bar{\delta}_{ij}\right]\right). \end{aligned} \tag{2.6}$$

The path ordering operator P can be dropped, because the expression in square brackets is independent of the ordering variable, E_k. Physically, the Sudakov operator represents the probability of no emission occurring between scales a and b.

All the elements in Eq. (2.1) are now defined. We can condense the algorithm using the amplitude operators $\mathbf{A}_n$, which satisfy the recurrence relation

$$\mathbf{A}_n(E) = \mathbf{V}_{E,E_n}\mathbf{D}_n^{\mu}\mathbf{A}_{n-1}(E_n)\mathbf{D}_{n\mu}^{\dagger}\mathbf{V}_{E,E_n}^{\dagger}\Theta(E \leq E_n), \tag{2.7}$$

where $\Theta(E \leq E_n)$ is a Heaviside step function that constraints the scale of each subsequent emission to be lower or equal than the last. The first element in the recursion is $\mathbf{A}_0 = \mathbf{H}$, and the last one will evolve down to $E = \mu$. Eq. (2.7) exposes the structure behind the algorithm: the amplitude operator at each step is related to the previous one by a real emission in the amplitude and conjugate amplitude, and subsequent evolution with virtual corrections at all orders down to the next scale. The amplitude operators $\mathbf{A}_n$ contain fixed order infrared poles, the cancellation of which only occurs after summing

over n and integrating over their phase-space.

The final step is to compute a general observable with

$$\Sigma(\mu) = \int \sum_n \mathrm{d}\sigma_n(\mu) u_n(\{k\}_n), \tag{2.8}$$

where u_n are the measurement functions that define the observable and depend on the set of momenta for all soft gluons $\{k\}_n$ at a given multiplicity n. In terms of the amplitude operators this is

$$\Sigma(\mu) = \int \sum_n \left(\prod_{i=1}^{n} \mathrm{d}\Pi_i \right) \mathrm{Tr}\, \mathbf{A}_n(\mu) u_n(\{k\}_n). \tag{2.9}$$

If we consider an observable that is fully inclusive below some scale ρ, the KLN theorem tells us that all real and virtual contributions will cancel below this scale, making it safe to set $\mu = \rho$. A detailed proof is given in [20], but it can be summarised by using the identities

$$\begin{aligned} \mathbf{V}^{\dagger}_{a,b}\mathbf{V}_{a,b} - 1 &= -\frac{\alpha_\mathrm{s}}{\pi} \int_a^b \frac{\mathrm{d}E}{E} \frac{\mathrm{d}\Omega_k}{4\pi} \mathbf{V}^{\dagger}_{E,b} \mathbf{D}^2(n_k) \mathbf{V}_{E,b}, \\ \frac{1}{2}\mathbf{D}^2(n_k) &= \sum_{i<j} (-\mathbf{T}_i \cdot \mathbf{T}_j) \omega_{ij}(n_k), \end{aligned} \tag{2.10}$$

where $\mathbf{D}^2(n_k) \equiv \mathbf{D}^{\mu}(n_k)\mathbf{D}_{\mu}(n_k)$, to rewrite the cross-section for n emissions as

$$\mathrm{d}\sigma_n = \left(\prod_{m=1}^{n} \mathrm{d}\Pi_m \right) \mathrm{Tr} \left(\mathbf{A}_n(E_n) - \int \mathrm{d}\Pi_{n+1} \mathbf{A}_{n+1}(E_{n+1}) \right). \tag{2.11}$$

The observable is then given by

$$\begin{aligned} \Sigma(\mu) = \mathrm{Tr}\, \mathbf{H} + \sum_{n=0}^{\infty} \int \left(\prod_{m=1}^{n+1} \mathrm{d}\Pi_m \right) \mathrm{Tr}(\mathbf{A}_{n+1}(E_{n+1})) \\ \times \left(u_{n+1}(\{k_1, ..., k_{n+1}\}) - u_n(\{k_1, ..., k_n\}) \right). \end{aligned} \tag{2.12}$$

The observable being fully inclusive below scale ρ implies that $u_{n+1}(\{k_1, ..., k_{n+1}\}) = u_n(\{k_1, ..., k_n\})$ for $E_{n+1} < \rho$, forcing all contributions to cancel. Therefore, for these types of observable we can set the evolution cutoff to ρ.

As mentioned before, it is possible to recast the soft evolution algorithm in Eq. (2.1) in a manifestly IR-finite form, tailored to the observable. Details can be found in [19, 20], but we can provide an overview. We consider an event-shape observable, where we divide the 4π solid angle into two regions, in gap and out of gap. The observable is inclusive over all out-of-gap radiation, and some constraint is set on the in-gap radiation. This is the gaps-between-jets observable we considered in Section 1.4.

We separate the soft-gluon anomalous dimension in two parts, $\mathbf{\Gamma} = \mathbf{\Gamma}_u + \overline{\mathbf{\Gamma}}_u$, where $\mathbf{\Gamma}_u$ contains virtual gluons in the out-of-gap region and $\overline{\mathbf{\Gamma}}_u$ in the in-gap region. We note that, if the observable is IRC-safe, $\overline{\mathbf{\Gamma}}_u$ should contain no infrared divergences. Then we express the Sudakov operator by exponentiating the in-gap $\overline{\mathbf{\Gamma}}_u$ and expanding in insertions of out-of-gap corrections $\mathbf{\Gamma}_u$. We then rewrite a recursive amplitude operator $\mathbf{A}'_n(E)$ as only resumming over the in-gap virtual gluons, and including at each step another out-of-gap real emission or out-of-gap virtual exchange. This amounts to an expansion in "non-globalness", where the global logarithms are captured by exponentiating $\overline{\mathbf{\Gamma}}_u$, which do not cancel against any real emissions, and the non-global corrections are perturbatively included by adding out-of-gap contributions. For each number of emissions n, the operators $\mathbf{A}'_n(E)$ are explicitly IR-finite.

For the remainder of this work we will focus on the general formulation provided at the beginning of this chapter, as we intend to capture all-orders non-global logarithms, and we are also going to sum over all real emissions and integrate over their phase-space. Analytically, this removes all infrared divergences, however computationally we will need to be careful to keep the simulation numerically stable. We leave this discussion for Section 3.1.6.

We end here the summary of the general soft-gluon evolution algorithm. In a sense, while fundamental and all-encompassing, what we have written so far is remarkably simple. A lot of the complexity in implementing amplitude-level evolution arises in the details of the implementation. We now proceed by delving into the most pressing question left unanswered in this section: how to represent the evolution in colour space.

2.1.1 Evolution in the colour flow basis

The general algorithm suffers from rapidly increasing complexity in its colour structures, which poses the greatest challenge to an exact colour treatment of large logarithmic contributions. We will use the colour flow basis to disentangle the evolution in colour space, and systematically break down all contributions to subleading colour. The fundamental colour mechanics described in what follows are at the heart of `CVolver`, and will be essential to the rest of this work.

The core principle behind this basis is to decompose the colour structure of the process into a set of 'flows' connecting each particle with colour to another with anti-colour. A first glimpse of this method was provided in Section 1.2. The colour flow description arises from the Fierz identity:

$$\begin{aligned}\sum_a t^a_{ij} t^a_{kl} &= \frac{1}{2}\left(\delta_{il}\delta_{kj} - \frac{1}{N_\mathrm{c}}\delta_{ij}\delta_{kl}\right) \\ &\equiv \frac{1}{2}\left(\begin{array}{cc} i & \longrightarrow \\ j & \longleftarrow \end{array} \begin{array}{c} l \\ k \end{array} \right) - \frac{1}{2N_\mathrm{c}}\left(\begin{array}{cc} i & \curvearrowright \\ j & \curvearrowleft \end{array} \begin{array}{c} l \\ k \end{array} \right)\end{aligned} \tag{2.13}$$

where $t^a_{ij} \in \mathfrak{su}(N_\mathrm{c})_{\text{fundamental}}$ are the generators of $\mathrm{SU}(N_\mathrm{c})$. The index a runs over the number of generators, the dimension of $\mathfrak{su}(N_\mathrm{c})_{\text{adjoint}}$, and i and j run over the number of colours, the dimension of $\mathfrak{su}(N_\mathrm{c})_{\text{fundamental}}$. The second line shows how the Kronecker deltas can be understood diagrammatically as "connecting" the fundamental colour indices of two quarks.

Eq. (2.13) separates the $\mathrm{SU}(N_\mathrm{c})$ propagator for a gluon i into a $\mathrm{U}(N_\mathrm{c})$ gluon propagator, with associated colour and anti-colour indices c_i and $\bar{c}_i$, corresponding to the first set of parallel flows, and a $\mathrm{U}(1)$ propagator, which does not carry any colour nor anti-colour indices, corresponding to the second colourless flow. We will refer to these colourless flows as singlet gluons[1].

Outgoing quarks carry colour, incoming quarks carry anti-colour, and vice versa for anti-quarks. Gluons carry both colour and anti-colour, and singlet gluons carry neither. The colour indices c_i are labelled from 1 to n_{lines}, the total number of colour lines, and are assigned to each of the colour-carrying partons. The anti-colour indices $\bar{c}_i$ are then

[1] Multiple names have been proposed for this unphysical particle, perhaps the most accurate being "traceless-condition gluon", since it is required to maintain the traceless property of the Lie algebra of $\mathrm{SU}(N_\mathrm{c})$, which distinguishes it from $\mathrm{U}(N_\mathrm{c})$. In this work we refer to them as singlet gluons for brevity.

assigned to each of the anti-colour carrying partons.

We show an example towards the left in Fig. 2.1. There are four outgoing partons: a quark, two gluons, and an anti-quark. These are labelled as partons 1 through 4. In total there are three colour lines and three anti-colour lines, which are labelled 1 through 3 and $\bar{1}$ through $\bar{3}$ respectively. Quark 1 has colour line 1, gluon 2 has colour line 2 and anti-colour line $\bar{1}$, gluon 3 has colour line 3 and anti-colour line $\bar{2}$, and lastly anti-quark 4 has anti-colour line $\bar{3}$. We refer to this association of colour and anti-colour line indices to partons as our "crossing".

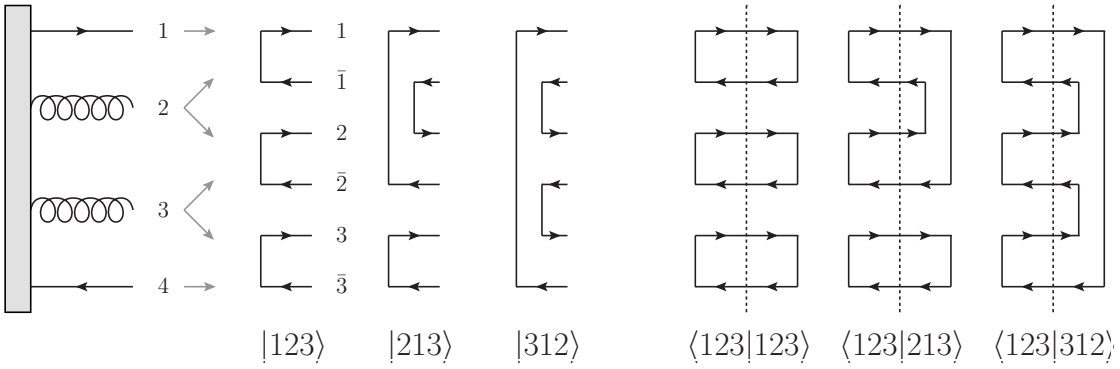


Figure 2.1: Diagrammatic representation of the colour flow basis applied to a process with external final-state legs $qgg\bar{q}$. From the left: The four legs and their labels. The grey arrows indicate which colour and anti-colour line indices belong to each leg. These indices are connected in different ways to form a colour flow state; three out of the six possible colour flows are shown. To the right, these colour flows are contracted and their inner product is given by N_c to the power of the number of closed loops formed, following from their Kronecker delta definition. Figure taken from [20].

The basis vectors consist of all possible ways to connect each colour index with an anti-colour one. Algebraically, they are a series of Kronecker deltas contracting each pair of connected indices:

$$|\sigma\rangle = \left| \begin{matrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{matrix} \right\rangle = \delta^{c_1}_{\sigma(c_1)} \cdots \delta^{c_n}_{\sigma(c_n)}, \tag{2.14}$$

where $\sigma(c_i)$ is the anti-colour index connected to c_i in state $|\sigma\rangle$. We have also simplified $n_\text{lines} \equiv n$ as it should be clear that in this context it refers to the number of colour lines, and not the number of emissions. We will refer to the basis vectors $|\sigma\rangle$ as colour flows, hinting at a physical interpretation of colour "flowing" through the system, and being conserved at each vertex.

For any given process, there are $n = n_g + n_q = n_g + n_{\bar{q}}$ colour lines, where n_g, n_q, and

$n_{\bar{q}}$ are the total number of external gluons, quarks and anti-quarks respectively, including hard process partons. This leads to $n!$ possible colour flow permutations. Each colour flow is naturally represented by one of these permutations, and we define the short-hand

$$|\sigma\rangle = \left|\begin{matrix} 1 & 2 & 3 \\ \bar{3} & \bar{1} & \bar{2} \end{matrix}\right\rangle \equiv |312\rangle \tag{2.15}$$

to efficiently label each flow as a permutation. Since the top row of colour indices is constant and ordered, we can omit it. Note we also drop the bars on the anti-colour line indices.

Three possible colour flows are shown in Fig 2.1. Diagrammatically, colour flows can be drawn by keeping the positions of each colour and anti-colour line fixed and connecting them in different ways. Three possible colour flow diagrams are drawn: $|123\rangle$, $|213\rangle$, and $|312\rangle$.

A closed colour loop results in a factor $\delta_{ii} = N_\mathrm{c}$. The scalar product of two colour flow basis tensors can be evaluated using their Kronecker delta definition:

$$\langle\sigma|\ \tau\rangle = \delta^{c_1}_{\sigma(c_1)} \dots\ \delta^{c_n}_{\sigma(c_n)}\ \delta^{\tau(c_1)}_{c_1} \dots\ \delta^{\tau(c_n)}_{c_n} = N_\mathrm{c}^{n-\mathrm{T}(\sigma,\tau)}, \tag{2.16}$$

where $\mathrm{T}(\sigma,\tau)$ is the minimum number of transpositions by which σ and τ differ. Diagrammatically, the number of closed loops formed will be the number of colour flows minus the number of transpositions, and the scalar product is $N_\mathrm{c}^{\#\text{ of loops}}$. Three example scalar products are shown on the right side of Fig. 2.1, and from the contracted colour flow diagrams we obtain $\langle 123|123\rangle = N_\mathrm{c}^3$, $\langle 123|213\rangle = N_\mathrm{c}^2$, and $\langle 123|312\rangle = N_\mathrm{c}$.

As we saw previously, each colour flow is represented by a permutation of n elements, which leads to $n!$ possible colour flow permutations. In Fig. 2.1 we have $n = 3$, thus there is a total of six basis vectors, of which three are shown. This basis is overcomplete, because $n!$ is larger than the actual number of possible colour states [101]. Amplitudes can be decomposed as $|\mathcal{A}\rangle = \sum_\sigma \mathcal{A}_\sigma |\sigma\rangle$, where σ labels the colour flow basis tensors, and the density matrices of colour operators relate to the basis vectors via

$$\mathbf{A} = \sum_{\sigma,\tau} \mathcal{A}_{\tau\sigma} |\tau\rangle\langle\sigma|. \tag{2.17}$$

However, the colour flow basis is non-orthonormal and overcomplete, so in order to calculate traces over operators in this basis we need to introduce a dual basis such that:

$$\langle\sigma|\,\tau] = [\sigma|\,\tau\rangle = \delta_{\sigma\tau}, \qquad \text{and} \qquad \sum_\sigma |\sigma\rangle\,[\sigma| = \sum_\sigma |\sigma]\,\langle\sigma| = \mathbf{1}. \tag{2.18}$$

Inserting the completeness relation, the trace of an operator in colour space becomes

$$\mathrm{Tr}\,\mathbf{A} = \mathrm{Tr}[\mathcal{A}\mathcal{S}] = \sum_{\sigma,\tau} [\tau|\,\mathbf{A}\,|\sigma]\,\langle\sigma|\,\tau\rangle\,, \tag{2.19}$$

where $\mathcal{A}$ has elements $\mathcal{A}_{\tau\sigma} = [\tau|\,\mathbf{A}_n\,|\sigma] = \langle\tau|\,\mathbf{A}_n\,|\sigma\rangle$ and the scalar product matrix has elements $S_{\tau\sigma} = \langle\sigma|\,\tau\rangle$. We note that an explicit representation of the dual basis is not needed for our purposes; the physical cross-section is built from matrix elements of colour operators and scalar products of colour flows, as shown in Eq. (2.19). This structure foreshadows Chapter 3, as we will need to keep track of factors of N_c from both the evolution operators $\mathcal{A}_{\tau\sigma}$ and the scalar product matrix $\mathcal{S}_{\tau\sigma}$. For now, we will focus on what the colour evolution looks like within the former, which contains recursive insertions of the colour operators $\mathbf{T}_i$ and $\exp\left[\sim \sum_{i\neq j} \mathbf{T}_i \cdot \mathbf{T}_j\right]$. This dual basis is discussed in more detail in [20] and [18], but for our purposes the results here are enough.

2.1.2 The colour charge operators

The two operators in colour space appearing in our soft gluon evolution algorithm are the colour charge operator $\mathbf{T}_i$, in the real emission operator, and the exponentiated colour charge products $\mathbf{T}_i \cdot \mathbf{T}_j$, in the Sudakov operator. The former can be decomposed as

$$\mathbf{T}_i = \lambda_i \mathbf{t}_{c_i} - \bar{\lambda}_i \bar{\mathbf{t}}_{\bar{c}_i} - \frac{1}{N_\mathrm{c}}(\lambda_i - \bar{\lambda}_i)\mathbf{s}, \tag{2.20}$$

where the operator $\mathbf{t}$ connects an additional gluon to the colour flow c_i, $\bar{\mathbf{t}}$ connects it to the anti-colour flow $\bar{c}_i$ and $\mathbf{s}$ emits a singlet gluon. The values of λ_i and $\bar{\lambda}_i$ for each particle can be found in Table 2.1. A singlet gluon has $\lambda_s = \bar{\lambda}_s = 0$, so they do not emit at all. Gluons carry $\lambda_g = \bar{\lambda}_g = 1/\sqrt{2}$, so they cannot emit singlets. Only quarks and anti-quarks can emit both gluons and singlets.

Parton i	λ_i	$\bar{\lambda}_i$
Quark	$\sqrt{T_R}$	0
Anti-quark	0	$\sqrt{T_R}$
Gluon	$\sqrt{T_R}$	$\sqrt{T_R}$
Singlet gluon	0	0

Table 2.1: The λ variables assigned to each different type of parton. T_R takes the conventional value in QCD of $1/2$. A non-zero λ_i indicates parton i carries colour, and a non-zero $\bar{\lambda}_i$ indicates it carries anti-colour. These values correspond to outgoing particles. For incoming particles the values of λ_i and $\bar{\lambda}_i$ are exchanged.

The gluon emission operators are algebraically defined

$$\mathbf{t}_{c_i} |\sigma\rangle = \mathbf{t}_{c_i} \left| \begin{matrix} 1 & \cdots & c_i & \cdots & n \\ \sigma(1) & \cdots & \sigma(c_i) & \cdots & \sigma(n) \end{matrix} \right\rangle \tag{2.21}$$

$$= \left| \begin{matrix} 1 & \cdots & c_i & \cdots & n & n+1 \\ \sigma(1) & \cdots & \overline{n+1} & \cdots & \sigma(n) & \sigma(c_i) \end{matrix} \right\rangle, \tag{2.22}$$

connecting the new gluon's colour line to anti-colour $\sigma(c_i)$, and reconnecting the colour line c_i to the new anti-colour line $\overline{n+1}$. The anti-colour line emission operators $\bar{\mathbf{t}}_{\sigma(c_i)}$ emit a gluon from an anti-colour line index $\sigma(c_i)$, which is equivalent to emitting from the colour index connected to it, c_i. Thus, $\bar{\mathbf{t}}_{\sigma(c_i)} = \mathbf{t}_{c_i}$.

The singlet emission operator is defined

$$\mathbf{s} |\sigma\rangle = \mathbf{s} \left| \begin{matrix} 1 & \cdots & n \\ \sigma(1) & \cdots & \sigma(n) \end{matrix} \right\rangle = \left| \begin{matrix} 1 & \cdots & n & n+1 \\ \sigma(1) & \cdots & \sigma(n) & \overline{n+1} \end{matrix} \right\rangle, \tag{2.23}$$

where a singlet is identified as a gluon whose colour and anti-colour indices are connected to each other in $|\sigma\rangle$. The operator $\mathbf{s}$ does not need a label as its action on $|\sigma\rangle$ is equivalent when emitted from any line. However, its coefficient in $\mathbf{T}_i$ still depends on the hard leg i, so that label cannot be dropped.

The action of the gluon emission operators is diagrammatically represented in Fig. 2.2. They create a new pair of colour and anti-colour lines which connect to the existing colour flow, in the case of $\mathbf{t}_{c_i}$ and $\bar{\mathbf{t}}_{\sigma(c_i)}$, or connect to themselves, independently of the existing colour flow, in the case of $\mathbf{s}$.

In the general soft evolution algorithm Eq. (2.1) the emission operators come in pairs

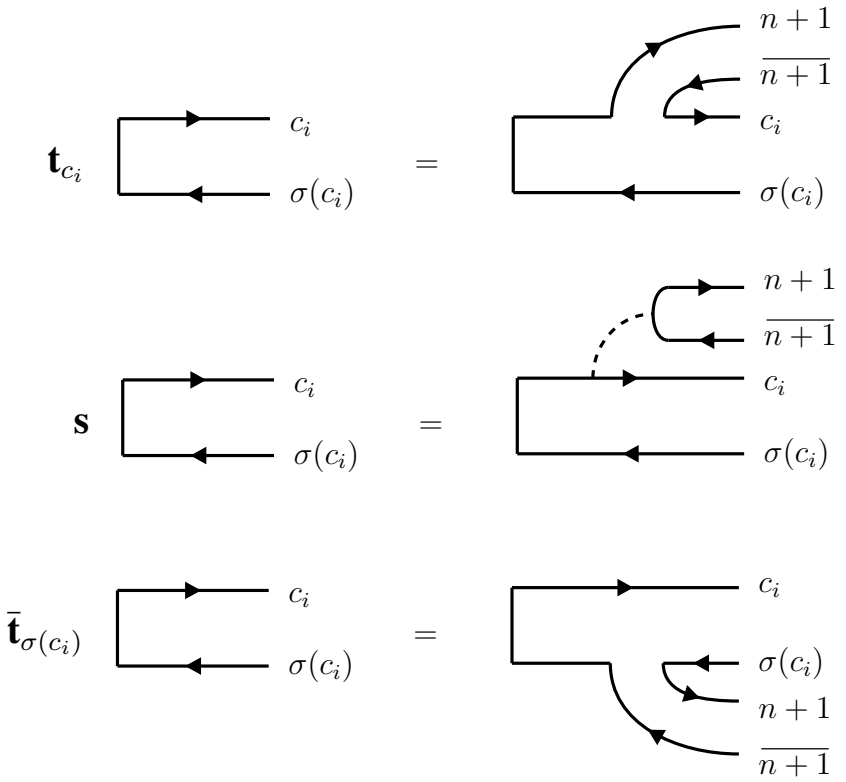


Figure 2.2: Diagrammatic representation of the emission operators $\mathbf{t}_{c_i}$, $\mathbf{s}$ and $\bar{\mathbf{t}}_{\sigma(c_i)}$. The first and last emit a gluon with colour and anti-colour, and reconnects these lines to the emitting colour connected pair. Their action is equivalent when c_i is colour connected to $\sigma(c_i)$. The $\mathbf{s}$ operator emits a singlet, which carries neither colour nor anti-colour, and is represented as a gluon colour connected to itself. It can be equivalently emitted from any line, as it does not interact with the existing colour flow.

$\mathbf{D}_i^\mu$ and $\mathbf{D}_{i\mu}^\dagger$, and act to the left and right respectively of the density operator. Their matrix elements can be contracted at cross-section level:

$$[\tau|\,\mathbf{D}_k^\mu\,|\sigma\rangle\,\langle\bar{\sigma}|\,\mathbf{D}_{k\mu}^\dagger\,|\bar{\tau}] = \sum_{i\neq j}[\tau|\,\mathbf{T}_i\,|\sigma\rangle\,\langle\bar{\sigma}|\,\mathbf{T}_j^\dagger\,|\bar{\tau}]\,\omega_{ij}(n_k), \tag{2.24}$$

which factorises into the antenna function from Eq. (2.5) and the matrix elements of the colour charge operator, acting on the amplitude and conjugate amplitude colour flows

respectively. We can write a complete algebraic expression for the colour operators [20]:

$$
\begin{aligned}
[\sigma|\mathbf{T}_i\mathbf{A}\mathbf{T}_j^\dagger|\tau] &= \sum_{\alpha,\beta} [\sigma|\,\mathbf{T}_i\,|\alpha\rangle\,\langle\beta|\,\mathbf{T}_j^\dagger\,|\tau]\,[\alpha|\,\mathbf{A}\,|\beta] \\
&= \Bigg\{ \Big(-\lambda_i\bar{\lambda}_j\delta_{c_i\sigma^{-1}(\bar{c}_n)}\delta_{\bar{c}_j\tau(c_n)} - (i,\sigma\leftrightarrow j,\tau)\Big) \\
&\quad + \lambda_i\lambda_j\delta_{c_i\sigma^{-1}(\bar{c}_n)}\delta_{c_j\tau^{-1}(\bar{c}_n)} + \bar{\lambda}_i\bar{\lambda}_j\delta_{\bar{c}_i\sigma(c_n)}\delta_{\bar{c}_j\tau(c_n)} \\
&\quad - \frac{1}{N_c}(\lambda_i\delta_{c_i\sigma^{-1}(\bar{c}_n)} - \bar{\lambda}_i\delta_{\bar{c}_i\sigma(c_n)})(\lambda_j - \bar{\lambda}_j)\delta_{c_n\tau^{-1}(\bar{c}_n)} \\
&\quad - (i,\sigma\leftrightarrow j,\tau) \\
&\quad + \frac{1}{N_c^2}(\lambda_i - \bar{\lambda}_i)(\lambda_j - \bar{\lambda}_j)\delta_{c_n\sigma^{-1}(\bar{c}_n)}\delta_{c_n\tau^{-1}(\bar{c}_n)}\Bigg\} \\
&\quad \times [\tau\backslash n|\mathbf{A}|\sigma\backslash n],
\end{aligned}
\tag{2.25}
$$

where $|\sigma\backslash n]$ is the flow $|\sigma]$ where c_n and $\bar{c}_n$ are removed, the inverse of Eqs. (2.21) and (2.23). The second line includes the contributions where one gluon is emitted off a colour line and the other off an anti-colour line. The third line includes contributions where both gluons are emitted off either two colour lines or two anti-colour lines. The fourth line corresponds to one gluon and one singlet emission, and the sixth line to two singlet emissions. This expression is unwieldy. In terms of subleading colour corrections, it is necessary to consider not only the explicit factors of $1/N_\mathrm{c}$ that have appeared here, but also what impact the emissions will have on the scalar product matrix in Eq. (2.19). We will return to it and break it down into palatable pieces in Section 2.3. We move on to consider the colour charge products, $\mathbf{T}_i \cdot \mathbf{T}_j$, which appear in the soft gluon anomalous dimension, Eq. (2.4). Applying the definition of the colour charge operators, Eq. (2.20), we end up with the products $\mathbf{t}_{c_i} \cdot \mathbf{t}_{c_j}$, $\mathbf{t}_{c_i} \cdot \bar{\mathbf{t}}_{\bar{c}_j}$, $\mathbf{t}_{c_i} \cdot \mathbf{s}$ and $\mathbf{s} \cdot \mathbf{s}$. These dot products are defined by applying both colour emission operators to their respective lines, and contracting the new colour and anti-colour indices.

We show diagrammatically in Fig. 2.3 the results for the dot products without singlets. Note that we start counting from zero for the colour and anti-colour line labels, which is a convention we will maintain for the rest of this work.[2] We identify three cases: exchange between two colour lines (top row), exchange between a colour line and its colour-connected anti-colour line (middle row), and exchange between a colour line and an anti-colour line which are not connected (bottom row). Other exchanges are equivalent to one

[2]This is for consistency with the `C++` implementation, discussed in Chapter 3.

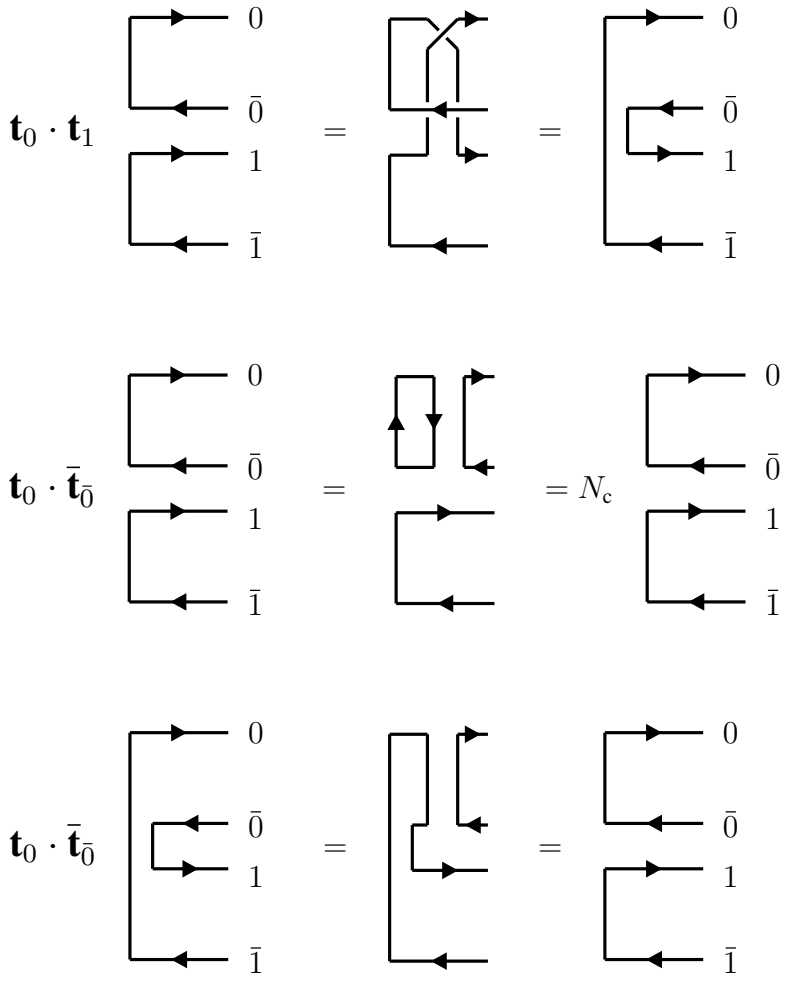


Figure 2.3: Diagrammatic representation of the $\mathbf{t}_0 \cdot \mathbf{t}_1$ operator (top row), the $\mathbf{t}_0 \cdot \bar{\mathbf{t}}_{\bar{0}}$ operator when 0 and $\bar{0}$ are colour connected (middle row) and the $\mathbf{t}_0 \cdot \bar{\mathbf{t}}_{\bar{0}}$ operator when 0 and $\bar{0}$ are not colour connected (bottom row). In the top and bottom cases a swap occurs between the initial and final colour flows. In the middle case the flow remains unchanged, and a factor of N_c is gained.

of these: in particular $\bar{\mathbf{t}}_{\sigma(c_i)} \cdot \bar{\mathbf{t}}_{\sigma(c_j)}$ will always be equivalent to $\mathbf{t}_{c_i} \cdot \mathbf{t}_{c_j}$. The top and bottom cases result in a rearrangement of the colour flow:

$$(\mathbf{t}_0 \cdot \mathbf{t}_1)|01\rangle = |10\rangle \qquad \text{and} \qquad (\mathbf{t}_0 \cdot \bar{\mathbf{t}}_{\bar{0}})|10\rangle = |01\rangle, \tag{2.26}$$

while the middle row leaves the colour flows intact, with an additional factor of N_c from the closed colour loop:

$$(\mathbf{t}_0 \cdot \bar{\mathbf{t}}_{\bar{0}})|01\rangle = N_c|01\rangle. \tag{2.27}$$

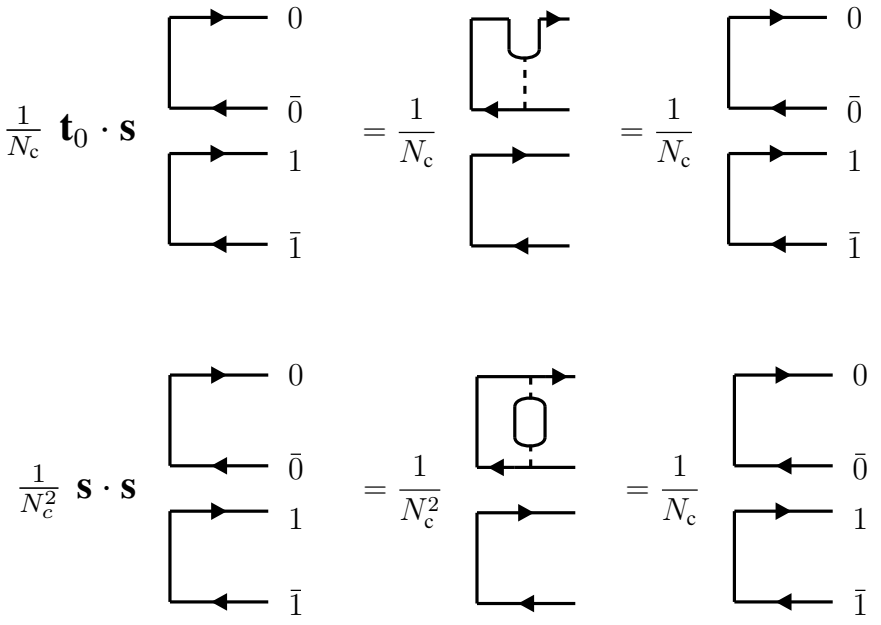


Figure 2.4: Diagrammatic representation of the $\mathbf{t}_0 \cdot \mathbf{s}$ operator (top row), and the $\mathbf{s} \cdot \mathbf{s}$ operator (bottom row). In both cases, the colour flows remain unchanged after applying the operators; their action is completely independent of which lines they connect to. Explicit factors of $1/N_c$ are included from the definition of $\mathbf{T}_i$, which in combination with the closed colour loop in $\mathbf{s} \cdot \mathbf{s}$, result in an overall factor of $1/N_c$ for both cases.

We will refer to a transposition of two elements in the colour flow permutation as a "swap", and this mechanism will be fundamental to the evolution in colour. Whether $\mathbf{t}_0 \cdot \bar{\mathbf{t}}_{\bar{0}}$ causes a swap or not depends on the colour flow it acts on, as it can either create a closed loop within a single colour connection, or rearrange two separate colour connections. However, $\mathbf{t}_0 \cdot \mathbf{t}_1$ always leads to a swap as two colour lines can never be connected to each other.

We now consider the exchanges which include an $\mathbf{s}$ operator, shown in Fig. 2.4. In all cases the flows remain unchanged:

$$\frac{1}{N_c}(\mathbf{t}_0 \cdot \mathbf{s})|01\rangle = \frac{1}{N_c^2}(\mathbf{s} \cdot \mathbf{s})|01\rangle = \frac{1}{N_c}|01\rangle. \tag{2.28}$$

and the action of these operators is completely independent of the colour flows they act on.

Similar to the real emissions, it is possible to put together all these cases to obtain an

algebraic expression for the colour charge product operator [20]:

$$\begin{aligned}[\tau\,|\mathbf{T}_i\cdot\mathbf{T}_j|\,\sigma\rangle = &-N_c\delta_{\tau\sigma}\left(\lambda_i\bar{\lambda}_j\delta_{c_i,\sigma^{-1}(\bar{c}_j)}+\lambda_j\bar{\lambda}_i\delta_{c_j,\sigma^{-1}(\bar{c}_i)}\right)\\ &-\frac{N_c\delta_{\tau\sigma}}{N_c^2}\left(\lambda_i-\bar{\lambda}_i\right)\left(\lambda_j-\bar{\lambda}_j\right)\\ &+\sum_{(ab)}\delta_{\tau_{(ab)},\sigma}\left(\lambda_i\lambda_j\delta_{(ab),(c_ic_j)}+\bar{\lambda}_i\bar{\lambda}_j\delta_{(ab),(\sigma^{-1}(\bar{c}_i)\sigma^{-1}(\bar{c}_j))}\right.\\ &\left.-\lambda_i\bar{\lambda}_j\delta_{(ab),(c_i,\sigma^{-1}(\bar{c}_j))}-\lambda_j\bar{\lambda}_i\delta_{(ab),(c_j,\sigma^{-1}(\bar{c}_i))}\right),\end{aligned} \tag{2.29}$$

where (ab) denotes an ordered pair of indices, and $|\tau_{(ab)}\rangle$ is the colour flow $|\tau\rangle$ with elements a and b swapped. The first line contains the colour-diagonal contributions (i.e., the colour flow remains unchanged) with an overall factor of N_c. These are the terms that include the middle row in Fig. 2.3. The third and fourth lines contain the off-diagonal contributions, with one swap between the initial and final colour flows, which correspond to the top and bottom rows in Fig. 2.3. Lastly, the second row contains the colour-diagonal contributions with an overall factor of $1/N_\mathrm{c}$, which include both rows in Fig. 2.4. Note that this matrix element is only non-zero when $|\sigma\rangle = |\tau\rangle$ or when they differ at most by one swap.

We have identified three types of contribution, and it is very useful to separate the anomalous dimension into three corresponding terms:

$$[\tau|\,\mathbf{\Gamma}\,|\sigma\rangle = \left(N_\mathrm{c}\Gamma_\sigma+\frac{1}{N_\mathrm{c}}\rho\right)\delta_{\tau\sigma}+\Sigma_{\tau\sigma} \tag{2.30}$$

where all the factors in the anomalous dimension have been absorbed, including the angular integral and Coulomb exchanges. Thus, these terms depend on the kinematic configuration of all legs, as well as the colour flows before and after the virtual exchange. The three terms are defined as follows:

- Γ_σ: the colour-diagonal, leading-colour contributions, arising from $\mathbf{t}_{c_i}\cdot\bar{\mathbf{t}}_{\bar{c}_j}$, where c_i and $\bar{c}_j$ are colour connected.
- ρ: the colour-diagonal, subleading colour contributions, arising from $\mathbf{t}_{c_i}\cdot\mathbf{s}$ and $\mathbf{s}\cdot\mathbf{s}$.
- $\Sigma_{\tau\sigma}$: the colour off-diagonal, subleading-colour contributions arising from $\mathbf{t}_{c_i}\cdot\mathbf{t}_{c_j}$ and $\mathbf{t}_{c_i}\cdot\bar{\mathbf{t}}_{\bar{c}_j}$, where c_i and $\bar{c}_j$ are not colour connected.

We have provided a breakdown of the leading and subleading colour effects that a single virtual gluon exchange can inflict on the evolution. This is all well and good, but the significant challenge of exponentiating the anomalous dimension remains. We will vanquish this problem in the next section.

2.2 The Sudakov operator

Exponentiating the soft anomalous dimension to obtain an expression for the Sudakov operator at all orders in colour was achieved in [18]. We will overview the derivation and explain the results, but the more rigorous algebraic details can be found there.

We start from the anomalous dimension as given in Eq. (2.30), broken down into three pieces according to their respective effect on the colour flow. The leading colour part, Γ_σ, is the identity in colour space, making it simple to exponentiate. From doing this we obtain $\exp(N_\mathrm{c}\Gamma_\sigma)$ as our zeroth-order approximation. We then treat ρ and $\Sigma_{\tau\sigma}$ as perturbations, where each insertion is accompanied by a factor of $1/N_\mathrm{c}$. We can write the full colour Sudakov operator as an infinite series in powers of $1/N_\mathrm{c}$, where each order includes another $\Sigma_{\tau\sigma}$ or ρ insertion. A matrix element of the Sudakov operator is then given by

$$\begin{aligned}[\tau|e^{\mathbf{\Gamma}}|\sigma\rangle = \delta_{\tau\sigma}R(\{\sigma\},\Gamma) + \sum_{l=1}^{d}\frac{(-1)^l}{N_\mathrm{c}^l}\sum_{k=0}^{l}\frac{(-\rho)^k}{k!} \\ \times \sum_{\{\sigma_0,\sigma_1,\cdots,\sigma_{l-k}\}}\delta_{\tau\sigma_0}\delta_{\sigma_{l-k}\sigma}\left(\prod_{\alpha=0}^{l-k-1}\Sigma_{\sigma_\alpha\sigma_{\alpha+1}}\right)R(\{\sigma_0,\sigma_1,\cdots,\sigma_{l-k}\},\Gamma),\end{aligned} \tag{2.31}$$

where $\Gamma = \{\Gamma_{\sigma_i} \mid \sigma_i \in \{\sigma_0,\sigma_1,\ldots,\sigma_{l-k}\}\}$, and we should pause here to properly unpack this expression. We have introduced the R functions, which are complex scalars, not matrix exponentials. They are in general functions of some sequence of colour flows $\{\sigma_0,\sigma_1,\cdots,\sigma_{l-k}\}$, and of the corresponding set of contributions $\{\Gamma_{\sigma_0},\Gamma_{\sigma_1},\ldots,\Gamma_{\sigma_{l-k}}\}$. The R functions are defined in closed form in [18]. For example, in the first term:

$$R(\{\sigma\},\Gamma) = e^{N_\mathrm{c}\Gamma_\sigma} \tag{2.32}$$

which is the zeroth-order approximation to the Sudakov operator we identified before.

The first summation is over powers of $1/N_\mathrm{c}$, up to some order d. We can take $d\to\infty$ for an exact Sudakov operator, or choose a finite d to set a maximum order in subleading colour accuracy. We will return to this in Chapter 3. The second summation is over how many ρ insertions are present at order $(1/N_\mathrm{c})^l$; from zero to l. The second line is more sophisticated, and includes the necessary number of $\Sigma_{\alpha\beta}$ insertions to reach $(1/N_\mathrm{c})^l$. As we saw before, $\Sigma_{\alpha\beta}$ is only non-zero when α and β differ by exactly one swap. For example,

if $l-k=2$, then all possible sequences $\Sigma_{\tau\alpha}\Sigma_{\alpha\sigma}$ need to be considered, where τ and σ are fixed colour flows, since we are evaluating $[\tau|e^{\mathbf{\Gamma}}|\sigma\rangle$, and α can be any colour flow which is exactly one swap away from τ and one swap away from σ. All such sequences of colour flows, of length $l-k$, are summed over by $\sum_{\{\sigma_0,\sigma_1,\cdots,\sigma_{l-k}\}}$, and the Kronecker deltas impose that the final and initial colour flows correspond with the matrix element. Then, the Σ contributions associated with each sequence of colour flows is inserted by $\left(\prod_{\alpha=0}^{l-k-1}\Sigma_{\sigma_\alpha\sigma_{\alpha+1}}\right)$. Finally, for each sequence in colour flows there is an accompanying R function.

We can look at the first few terms to further appreciate the structure of the series. The Sudakov operator, up to next-to-next-to-leading-colour (NNLC), is given by

$$\begin{aligned}[\tau|e^{\mathbf{\Gamma}}|\sigma\rangle = \delta_{\tau\sigma}R(\{\sigma\},\{\Gamma_\sigma\}) + \delta_{\tau\sigma}\frac{\rho}{N_\mathrm{c}}e^{-N_\mathrm{c}\Gamma_\sigma} - \frac{\Sigma_{\tau\sigma}}{N_\mathrm{c}}R(\{\tau,\sigma\},\{\Gamma_\tau,\Gamma_\sigma\}) \\ + \frac{1}{N_\mathrm{c}^2}\Bigg(\sum_{\{\tau,\alpha,\sigma\}}\Sigma_{\tau\alpha}\Sigma_{\alpha\sigma}R(\{\tau,\alpha,\sigma\},\{\Gamma_\tau,\Gamma_\alpha,\Gamma_\sigma\}) \\ - \rho\Sigma_{\tau\sigma}R(\{\tau,\sigma\},\{\Gamma_\tau,\Gamma_\sigma\}) + \frac{\rho^2}{2}\delta_{\tau\sigma}e^{-N_\mathrm{c}\Gamma_\sigma}\Bigg),\end{aligned} \tag{2.33}$$

where the first line contains the zeroth and first orders. The first order contains either a single ρ or single $\Sigma_{\tau\sigma}$ insertion. The second order is in lines two and three, suppressed by a factor $1/N_\mathrm{c}^2$, and contains three terms: a double Σ, double ρ, or a Σ and a ρ. In the first case, it is necessary to sum over all possible colour flow sequences $\{\tau,\alpha,\sigma\}$, such that τ is one swap away from α, which is itself one swap away from σ. Depending on the number of colour lines, this sum can grow significantly large. The R functions that appear here are given by

$$\begin{aligned}R(\{\tau,\sigma\},\Gamma) &= \frac{e^{-N_\mathrm{c}\Gamma_\tau}-e^{-N_\mathrm{c}\Gamma_\sigma}}{\Gamma_\tau-\Gamma_\sigma} \\ R(\{\tau,\alpha,\sigma\},\Gamma) &= \frac{e^{-N_\mathrm{c}\Gamma_\tau}}{(\Gamma_\tau-\Gamma_\alpha)(\Gamma_\tau-\Gamma_\sigma)} + (\tau\leftrightarrow\sigma) + (\tau\leftrightarrow\alpha) \\ R(\{\tau,\alpha,\tau\},\Gamma) &= -N_\mathrm{c}\frac{e^{-N_\mathrm{c}\Gamma_\tau}}{\Gamma_\tau-\Gamma_\alpha} + \frac{e^{-N_\mathrm{c}\Gamma_\alpha}-e^{-N_\mathrm{c}\Gamma_\tau}}{(\Gamma_\tau-\Gamma_\alpha)^2},\end{aligned} \tag{2.34}$$

where the last R function is for the case where τ swaps to α and then performs the same swap back to τ again. It it found by simply taking the limit $\sigma\to\tau$ for $R(\{\tau,\alpha,\sigma\},\Gamma)$.

We will discuss how the sums over sequences of colour flows and their corresponding R functions are implemented in Chapter 3.

The ρ contributions, although subleading in colour, are colour diagonal. It follows that we can easily exponentiate them to all orders, and expand only in the number of $\Sigma_{\tau\sigma}$ insertions. The Sudakov operator with resummed virtual singlet exchanges is given by

$$\begin{aligned}[\tau|\, e^{\mathbf{\Gamma}}\, |\sigma\rangle &= \delta_{\tau\sigma} R(\{\sigma\}, \{\Gamma'_\sigma\}) \\ &\quad + \sum_{l=1}^{d} \left(-\frac{1}{N_c}\right)^l \sum_{\{\sigma_0,\ldots,\sigma_l\}} \delta_{\tau\sigma_0}\delta_{\sigma_l,\sigma} \left(\prod_{\alpha=0}^{l-1} \Sigma_{\sigma_{\alpha+1},\sigma_\alpha}\right) R(\{\sigma_0,\ldots,\sigma_l\}, \Gamma')\end{aligned} \tag{2.35}$$

where $\Gamma'_\sigma = \Gamma_\sigma - \rho/N_c^2$, and $\Gamma' = \{\Gamma'_{\sigma_i} \mid \sigma_i \in \{\sigma_0, \sigma_1, \ldots, \sigma_l\}\}$. When d is set to zero, we refer to the evolution with resummed virtual singlets as the LC$'$ approximation, where the prime indicates we include some subleading colour terms beyond LC. To completely capture the Sudakov at some subleading order in colour we will need $d > 0$, and that is our N^{d}LC$'$ approximation. The LC$'$ approximation is equivalent to the replacement $N_c/2 \to C_F$ in the diagonal part of the virtual evolution, and we will study how well it captures subleading colour effects in Chapter 4.

For now, let us provide a summary of what we have achieved thus far. First, we used the colour flow basis to find representations of the colour structures that appear in the general soft evolution algorithm. We decomposed both the real emissions and anomalous dimension in terms of the possible ways they can guide the colour evolution. We then used the decomposition of the anomalous dimension as a starting point to exponentiate it, resumming all diagonal terms and inserting off-diagonal contributions as perturbations. In doing so, we have gained insight into the subleading colour structure of the Sudakov. On the other hand, for the real emissions we stopped at a somewhat-ugly algebraic expression in Eq. (2.25), without much understanding of its subleading colour structure or how it relates to the scalar product matrix (Eq. (2.19)). In order to rectify this, we need to take a short detour into the topic of rings and strings.

2.3 Rings and strings

The rings and strings formalism was introduced in [22], and it constitutes a set of basis functions for the kinematic part of the amplitudes described in Section 2.1. This formalism builds on the work by Dokshitzer and Marchesini on the "5th form factor" [102]. In the following we will outline the results, and explain how these structures emerge in the evolution of colour flows.

It is a well-known result that collinear poles can only emerge in colour-diagonal structures, [22] provides a short proof. Consequently, any colour sub-leading contributions must be free of collinear poles. This is not apparent in the formulation of the general algorithm in Section 2.1, and it seems almost unlikely due to the appearance of the dipole functions $\omega_{ij}(n_k)$, defined in Eq. (2.5), which contain poles as emission k becomes collinear with particles i or j.

These dipole functions are kinematic factors associated with both the virtual evolution, since they appear in the Sudakov operator $\mathbf{V}_{a,b}$, and the cross-section level contraction of the emission operators $\mathbf{D}_k^\mu$, given in Eq. (2.24). As discussed in Section 2.1, the soft evolution algorithm is IR-safe after summing over all emissions and integrating over their phase-space. However, in a Monte Carlo implementation, observables are evaluated by sampling random trajectories in kinematics and colour, and therefore a single event may contain uncancelled collinear poles. To circumvent this problem it is possible to implement a collinear cutoff which removes collinear poles from the reals and virtuals, keeping the cross-section intact due to the KLN theorem. This is entirely an implementation-side issue, and will be discussed in depth in Section 3.1.6.

Although we have a method for handling collinear divergences in the Monte Carlo, computationally their cancellation is still a major source of numerical instability. This creates plenty of incentive to develop an alternative set of basis kinematic functions which are explicitly regular in colour off-diagonal contributions, significantly reducing the computational complexity of cancelling all collinear enhancements. This is achieved with the ring and string basis functions, a name that will soon become self-explanatory. The shortest ring is given by

$$R_{k,l}^{i,j} = \omega_{ij}(q_n) - \omega_{ik}(q_n) - \omega_{jl}(q_n) + \omega_{kl}(q_n), \tag{2.36}$$

where $i, j, k, l \in \{1, 2, \ldots, n-1\}$ are all distinct particles, and q_n is the momentum of the n^{th} emission. The pole cancellation pattern is clear: ω_{ij} introduces collinear poles

associated with partons i and j, which are then subtracted by $\omega_{ik}(q_n)$ and $\omega_{jl}(q_n)$. These, in turn, introduce collinear poles along the directions of k and l, which are ultimately cancelled by adding $\omega_{kl}(q_n)$. A similar logic applies starting from any of the four dipole functions. This circular pattern is reminiscent of a ring, and one can conclude that a linear combination of any number of dipoles will be finite as long as it follows this closed ring structure. Diagrammatic expressions are shown in [22].

It follows that linear combinations of dipoles which do not form closed rings will contain uncancelled poles. These are referred to as strings, with the simplest one being:

$$S_i^{j,k} = \frac{1}{2}\Big(\omega_{ij}(q_n) + \omega_{ik}(q_n) - \omega_{jk}(q_n)\Big), \tag{2.37}$$

where again $i, j, k \in \{1, 2, \ldots, n-1\}$ are all distinct particles. There is a remaining collinear pole with respect to i. More details on the construction of this basis and its diagrammatic representation are in [22], however we only need the general forms in order to exploit the collinear safety of subleading colour evolution.

2.3.1 Colour evolution in terms of dipoles, rings and strings

We will consider the virtuals first. In Section 2.2 we stated an expansion of the Sudakov operator in powers of $1/N_c$. The kinematic factors in the colour sub-leading terms are contained in $\Sigma_{\tau\sigma}$, which was introduced in Eq. (2.30) as the part of the anomalous dimension operator responsible for colour off-diagonal contributions. An analytic expression for $\Sigma_{\tau\sigma}$ can be found by keeping track of the kinematic factors from Eq. (2.4) associated with the terms in the colour product operator that result in a swap of colour connections. This leads to

$$\Sigma_{\tau\sigma} = \sum_{i,l \text{ c.c. in } \sigma} \sum_{j,k \text{ c.c. in } \tau} \Omega_{ij} + \Omega_{kl} - \Omega_{ik} - \Omega_{jl}, \tag{2.38}$$

$$\Omega_{ij} = \frac{\alpha_s}{\pi} \ln\left(\frac{b}{a}\right) \int \frac{\mathrm{d}\Omega_n}{4\pi} \omega_{ij}(q_n), \tag{2.39}$$

where "i, l c.c. in σ" indicates that partons i and l are colour connected in the colour flow σ. We do not consider Coulomb gluons in this discussion, however the general algorithm and the Monte Carlo implementation are able to include them. We can now identify this as a ring, or more accurately, a solid angle integral over a ring, which can be evaluated

analytically to give

$$\int \frac{\mathrm{d}\Omega}{4\pi} R^{i,j}_{k,l} = 2 \ln \left(\frac{n_i \cdot n_j \;\; n_k \cdot n_l}{n_i \cdot n_k \;\; n_j \cdot n_l} \right), \tag{2.40}$$

which contains no poles. Since we work directly with factors of $\Sigma_{\tau\sigma}$ in the Sudakov expansion, we are already explicitly collinear safe in the colour-subleading evolution, and so there are no improvements to be made on the side of the virtuals.

The story is very different with respect to the real emissions. In a Monte Carlo algorithm it is beneficial to tailor the sampling of the evolution according to the matrix elements, in order to reduce the spread of event weights and therefore the statistics required. This is called importance sampling. Looking at the form of the real emission matrix elements in Eq. (2.24), it seems the obvious choice is to select an emitting pair i and j and sample the direction of the emission according to ω_{ij}. We can then select the colour flows after the emission from the list of accessible colour flows after the action of the $\mathbf{T}_i$ and $\mathbf{T}_j^\dagger$ operators.

This method, although straightforward, introduces two collinear poles with respect to the directions i and j. However, we know that if the contribution is subleading in colour, these poles should cancel. This prompts us to instead consider the possible trajectories a real emission can follow in colour space, and combine together the dipoles that contribute to the same trajectory.

We show the different possibilities in Fig. 2.5. We do not consider the emission of singlets for now. As we saw in Fig. 2.2, the action of $\mathbf{t}_{c_i}$ is equivalent to $\bar{\mathbf{t}}_{\sigma(c_i)}$, so we only need to focus on which colour-connected lines the emission is from. In the bottom row, and for each of the three diagrams, a gluon emission, depicted by wavy lines, is emitted in the amplitude (gluon emitted towards the left) and conjugate amplitude (gluon emitted towards the right). The left case corresponds to the gluon being emitted from the same colour connected lines on both sides, c_i connected to $\bar{c}_j$. The middle case has one colour line in common, c_i, but they differ in anti-colour line: $\bar{c}_j$ for the amplitude, and $\bar{c}_k$ for the conjugate amplitude. The right case corresponds to the gluons being emitted from entirely different colour connected lines: c_i to $\bar{c}_k$ in the amplitude, c_j to $\bar{c}_l$ in the conjugate amplitude. The top row shows the flows before the emission for each case, and the coloured dots indicate all the possible dipoles from which the emission can occur, such that the new colour and anti-colour lines attach as drawn in the bottom row. We categorise the three cases:

- Left: Emission off the same colour-connected pair in both the amplitude and conju-

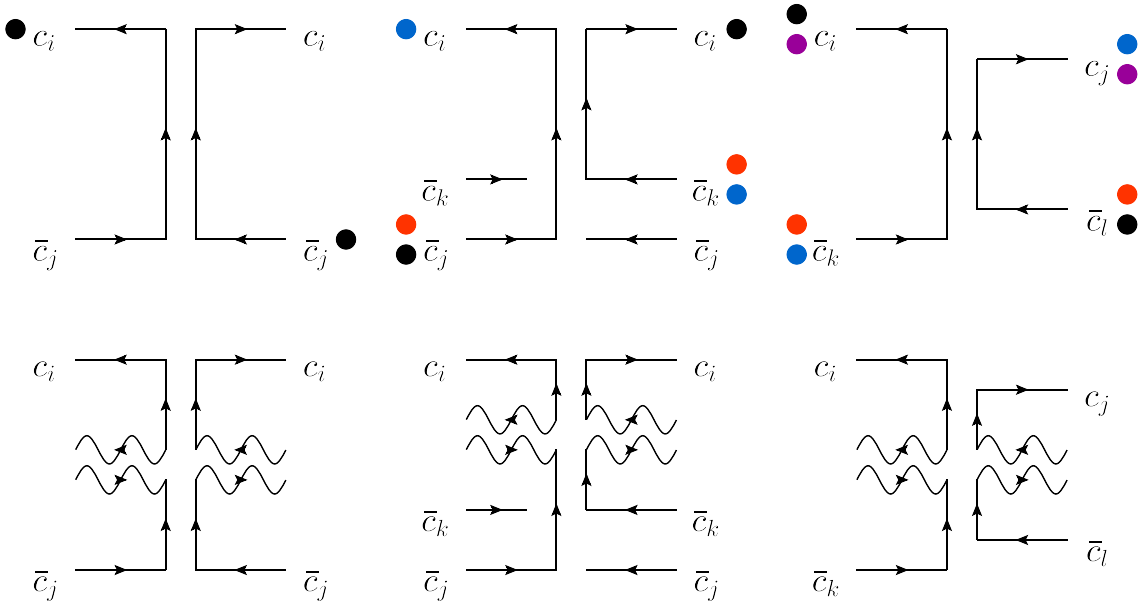


Figure 2.5: Top row of each diagram shows the colour flows before emission, the bottom row shows them after emission. Each diagram has the amplitude on the left, and the conjugate amplitude on the right. Each pair of lines labelled with the same coloured dot represents a possible dipole pair to emit from to obtain the flows directly below. Flows that do not participate in the emission are not drawn. The left diagram corresponds to a dipole emission, the middle one to a string, and the right one to a ring. Figure taken from [22].

gate amplitude. The associated kinematic factor is $-\omega_{ij}$. This is our dipole contribution.

- Middle: Emission off colour connections that only share either a colour line, or an anti-colour line. The associated kinematic factor is $-(\omega_{ij} + \omega_{ik} - \omega_{jk})$. This is our string contribution.

- Right: Emissions do not share colour lines or anti-colour lines in the amplitude and conjugate amplitude. The associated kinematic factor is $-(\omega_{il} + \omega_{kj} - \omega_{kl} - \omega_{ij})$. This is our ring contribution.

We should look at whether these are leading or subleading colour contributions. This is not a trivial task; factors of $1/N_c$ can come from both the real emission matrix elements themselves, but also the scalar product matrix, which appeared in Eq. (2.19). We can combine both effects and define the quantity

$$\xi_{ij}^{(\tau\sigma\bar{\sigma}\bar{\tau})} = \frac{\langle\tau|\,\bar{\tau}\rangle}{\langle\sigma|\,\bar{\sigma}\rangle}\left|[\tau|\,\mathbf{T}_i\,|\sigma\rangle\,\langle\bar{\sigma}|\,\mathbf{T}_j^\dagger\,|\bar{\tau}]\right|, \tag{2.41}$$

where $\{\sigma, \bar{\sigma}\}$ are the colour flows before the emission, in the amplitude and conjugate amplitude respectively, and $\{\tau, \bar{\tau}\}$ are the colour flows after the emission. The colour charge operator matrix elements will include any explicit factors of $1/N_c$, as well as become zero if the flows after the emission are not accessible from the initial flows. The $\langle\tau|\ \bar{\tau}\rangle\ /\ \langle\sigma|\ \bar{\sigma}\rangle$ fraction contains factors of N_c arising from the trajectory within the scalar product matrix. The quantity $\xi_{ij}^{(\tau\sigma\bar{\sigma}\bar{\tau})}$ is crucial for keeping track of the colour suppression inflicted by the real emissions, and it is also used for sampling trajectories in colour space in the Monte Carlo. Both of these points will be discussed in Sections 3.1.8 and 3.1.9. Let us go back to the three scenarios:

- Dipole emission: $\xi_{ij}^{(\tau\sigma\bar{\sigma}\bar{\tau})} = N_c$, always. The original closed colour loop becomes two loops. This is a Γ contribution.
- String emission: $\xi_{ij}^{(\tau\sigma\bar{\sigma}\bar{\tau})} = N_c$, always. Again, the original closed colour loop becomes two loops. However, a string emission can only occur if $|\sigma\rangle \neq |\bar{\sigma}\rangle$, and so we are off-diagonal in colour. This is a Σ contribution.
- Ring emission: $\xi_{ij}^{(\tau\sigma\bar{\sigma}\bar{\tau})} = N_c$ or $1/N_c$, depending on $|\sigma\rangle$ and $|\bar{\sigma}\rangle$. In the former case, a single closed loop becomes two. In the latter, two disjointed loops become one. In either case, this is a Σ contribution.

We have used the idea behind the pieces of the anomalous dimension in Eq. (2.30) to also categorise the real emissions. As before, a Γ contribution is leading colour and colour diagonal, while a Σ contribution is off-diagonal in colour. Note that ring contributions, which contain no collinear poles, are subleading in colour.

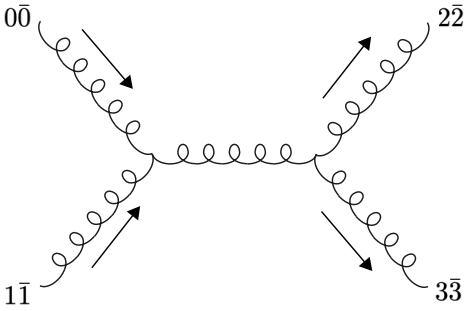
We need to consider also emissions of singlet gluons. These will be the ρ contributions. In terms of the contributing kinematic dipoles, these are more complicated as a singlet can be emitted off any quark, and the effect on the colour flows will be equivalent in all cases. If a singlet is emitted on one side and a gluon on the other, the associated kinematic factor is $\sum_{j\in q}\omega_{ij} - \sum_{l\in\bar{q}}\omega_{il} - \sum_{j\in q}\omega_{kj} + \sum_{l\in\bar{q}}\omega_{kl}$, where the gluon is emitted from colour connected legs i and k, and $\omega_{ii} = 0$. If two singlets are emitted, then we have to sum over all quarks and anti-quarks in both the amplitude and conjugate amplitude: $\sum_{i\in q}\left(\sum_{j\in q, i\neq j}\omega_{ij} - \sum_{l\in\bar{q}}\omega_{il}\right) + \sum_{k\in\bar{q}}\left(\sum_{j\in q}\omega_{kj} - \sum_{l\in\bar{q}, k\neq l}\omega_{kl}\right)$.

We should clarify our definition of a ring or string emission. We introduced them as in [22], such that a string contains one collinear divergence and a ring contains none. Then, we discussed how ring and string structures emerge when considering the Σ contributions to colour evolution, in Fig. 2.5. Notably, rings can only emerge if there are at least

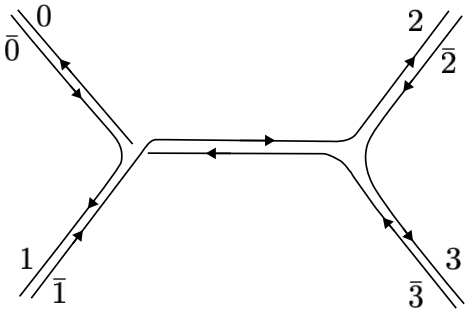
four distinct partons available to emit, and strings require three. This corresponds to the colour line indices in Fig. 2.5 belonging to a different particle each. However, as gluons carry colour and anti-colour, it is perfectly possible for two of the indices, for example c_i and $\bar{c}_l$, to belong to the same gluon. Then the kinematic factor of a ring collapses to $-(\omega_{kj} - \omega_{ki} - \omega_{ij})$. We will still refer to this as a ring emission, as we are more interested in its trajectory in colour space, and these types of emission fall into the category of off-diagonal, Σ contributions. Therefore, we only require two colour lines for a ring emission, and three for a string. Similarly, the $\Sigma_{\tau\sigma}$ structures (Eq. (2.38)) that appear in the anomalous dimension also collapse into fewer than four dipole functions if multiple colour line indices belong to the same parton.

Fig 2.6 provides a review of the colour flow basis and examples of emissions from a dipole, string, or ring in the $gg \to gg$ process. Fig. 2.6a shows the diagram, which is an s-channel gluon exchange, and the colour and anti-colour line labels of each hard leg. We will present results for this process in Chapter 4. Fig. 2.6b shows a possible colour flow contributing to the hard process matrix. Fig. 2.6c shows this flow using our standard diagrammatic representation of the permutation, and Fig. 2.6d shows an interference element in the hard density matrix, with the permutation of the amplitude to the left and the conjugate amplitude to the right. Lastly, Fig. 2.6e shows the resulting colour flows after a dipole, string or ring emission. The ring emission provides an example of what we discussed before; the ω_{11} dipole is zero, because colour line 1 and anti-colour line $\bar{1}$ belong to the same gluon.

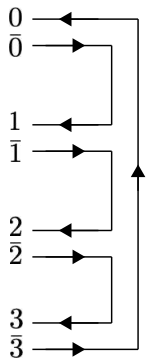
At this stage, we have presented the necessary groundwork to develop an amplitude-level Monte Carlo. The colour flow basis, its diagrammatic representation, and the action of the colour charge and colour product operators was defined in Section 2.1.1. We found that the anomalous dimension can be separated into three pieces: a leading colour part Γ, an off-diagonal subleading colour part, Σ, and a diagonal subleading colour part, ρ. We then used these pieces to find an expression for the Sudakov as an infinite series of subleading insertions in Section 2.2. Finally, we went back to the reals, and through the construct of rings and strings we categorised real emissions into the same three pieces: Γ, Σ, and ρ. At this stage we are ready to implement the evolution as a Monte Carlo algorithm, a task which is anything but trivial. This is the goal of Chapter 3.



(a) The hard process diagram corresponding to s-channel gluon exchange.

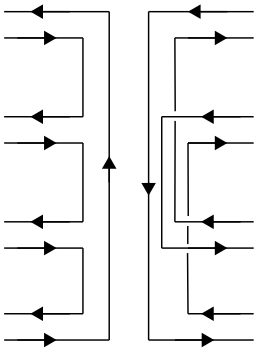


(b) The diagram of a possible configuration in the colour flow basis.



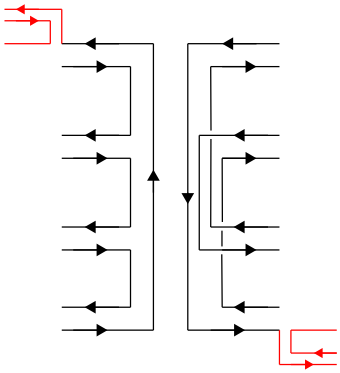
$|3012\rangle$

(c) The corresponding colour flow diagram.



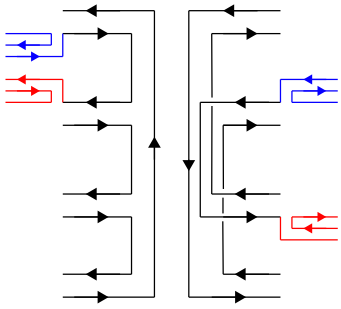
$|3012\rangle\langle 3201|$

(d) An interference contribution to the hard scattering density matrix.



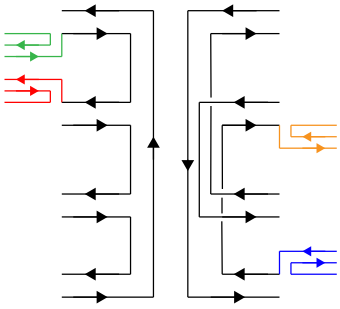
$|40123\rangle\langle 42013|$

Dipole: ω_{03}



$|34120\rangle\langle 34012|$

String: $\omega_{01} + \omega_{12} - \omega_{02}$



$|34120\rangle\langle 32041|$

Ring: $\omega_{03} + \omega_{11} - \omega_{01} - \omega_{13}$

(e) The emission of a gluon via the **D** operator from a dipole, a string and a ring. The kinematic factor and updated density matrix after the emission are also indicated. For the string and ring, only one gluon is emitted from a colour line in the amplitude and conjugate. The different possibilities are coloured differently.

Figure 2.6: An example to illustrate the colour flow basis and how we classify real emissions in terms of dipoles, strings and rings. Figure taken from [103].

Chapter 3

`CVolver`

`CVolver` is a framework for solving colour evolution equations of the class outlined in [19, 20]. It utilizes a plugin structure to solve structurally similar evolution equations, which might, however, have different physics applications or different levels of sophistication. In its current version, `CVolver` systematically expands around the large-N_c limit and combines techniques from normal parton showers as well as novel sampling approaches.

The framework originates from the development centred around [18], and was subsequently developed into the current structure by Matthew De Angelis, Jeff Forshaw, Patrick Kirchgaesser, Simon Plätzer and Fernando Torre González, with its first application to the studies presented in [23]. The first step in setting up `CVolver` was to address the virtual evolution, one of the major bottlenecks of soft gluon evolution algorithms. For this it is crucial to understand the inner workings of the colour flow basis and how it relates to the structure of anomalous dimension matrices. This needed to be complemented by a sampling of real-emission colour flows, along with a connection between the real and virtual weights, currently optimised to achieve almost constant weights for the real emission. The colour structures of gluon exchanges and emissions are now also available at two loops and for double emissions [104]. It was always important to keep track of different orders in N_c for the final cross section [20], a functionality which has been driving `CVolver` development since its beginning. Besides the soft gluon plugin, which has been built with the philosophy of the highest level of theoretical control for a specific observable, other plugins and interfaces are in development paving the way to a full event generator or other dedicated applications.

The present chapter summarizes the key ingredients of the soft gluon plugin [23, 105] and several improvements and additional data and analysis handling which have been developed in the context of this thesis to efficiently and flexibly simulate the full colour evolution of jet processes.

3.1 The Monte Carlo algorithm

We start by providing an overview of the whole algorithm, listing every step within one event. Then, we will describe in detail what happens at each step, what sampling distributions we use, and how the matrix elements are calculated, in the order followed by the algorithm. And to reinforce the theory with practical application, we will end the section by revisiting the overarching algorithm and printing out the calculation and results for a randomly selected event.

We can insert the colour flow basis completeness relation at each step in the general

evolution, given in Eq. (2.1), so that the evolved density matrix becomes a product of matrix elements of the real emission and Sudakov operators. The matrix elements contain sums over colour charge operators, in the case of the reals, and exponentiated colour product operators, in the case of the virtuals, and they each map a given colour flow vector to another. We rewrote the trace of the density matrix as a sum over final state colour flows in Eq. (2.19), and with these ingredients we are in good shape to implement the evolution using a Monte Carlo.

To calculate the cross-section, we simply need to integrate over the real emission phase-space elements, in Eq. (2.3), and sum over all possible trajectories in colour space. At each insertion of a real or virtual operator we choose one of the accessible colour flows after the operation. Summing over all possible trajectories is equivalent to taking the trace.

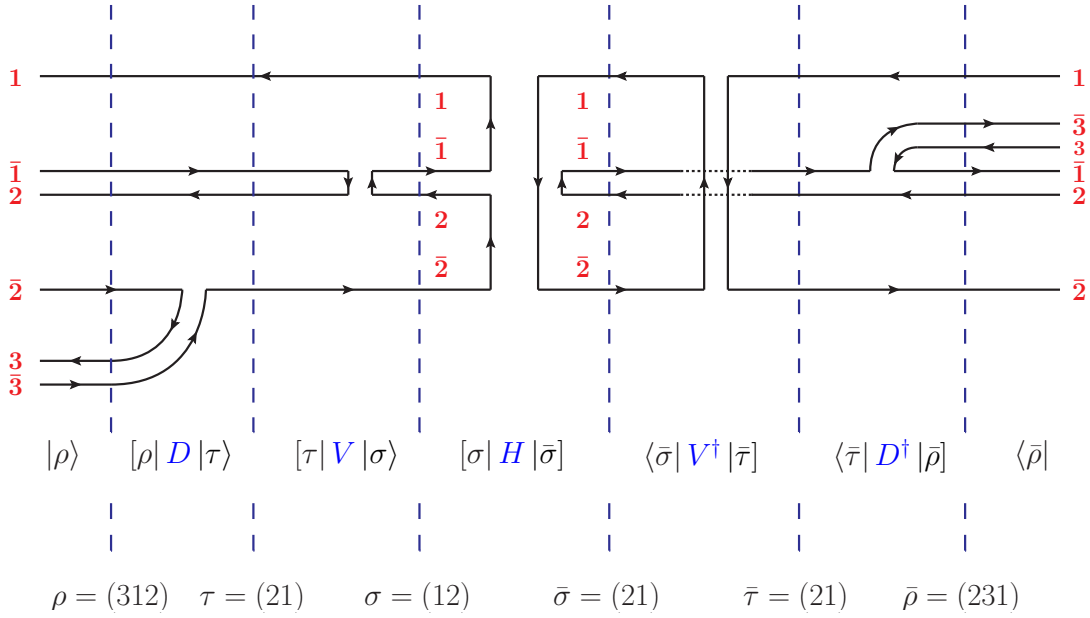


Figure 3.1: One contribution to the evolution operator up to the first real emission, starting from $|12\rangle\,[12|\,\mathbf{H}\,|21]\,\langle 21|$. The colour and anti-colour line indices are shown in red. The dotted lines indicate the colour flow state at each step. Figure taken from [23].

A diagrammatic example of one path in colour is shown in Fig. 3.1. The hard process has two colour and anti-colour lines each, which could be a contribution to, for example, $H \to gg$ or $q\bar{q} \to q\bar{q}$. First, the hard process colour states $\{\sigma, \bar{\sigma}\}$ are selected from the list of possible ones. Then, the scale of the first emission E_1 is sampled, and the new colour flow states $\{\tau, \bar{\tau}\}$ after virtual evolution are chosen from the list of accessible ones. The colour flows after the real emission $\{\rho, \bar{\rho}\}$ are then selected. The elements of the hard process, Sudakov operator, real emission operator, and their conjugates are multiplied together, to arrive at the complete matrix element. The process is repeated until one of the termination conditions is reached, which will be discussed later. This constitutes one event

in the Monte Carlo simulation, and the total cross-section is found by averaging over the weights of a large number of such events.

The iterative procedure c1an be expressed as a recursion relation, which encodes the evolution of the colour flow states at each step of the parton shower. Thus, we write the recursion from the general soft evolution algorithm, Eq. (2.7), in terms of matrix elements:

$$M_{\rho\bar{\rho}}(E) = -\frac{\alpha_s}{\pi}\frac{\mathrm{d}E}{E}\frac{\mathrm{d}\Omega}{4\pi}\sum_{\tau,\sigma,\bar{\tau},\bar{\sigma}}\left[\rho\right|\mathbf{D}_E\left|\tau\right\rangle\left[\tau\right|\mathbf{V}_{E,E'}\left|\sigma\right\rangle M_{\sigma\bar{\sigma}}(E')\left\langle\bar{\sigma}\right|\mathbf{V}^{\dagger}_{E,E'}\left|\bar{\tau}\right]\left\langle\bar{\tau}\right|\mathbf{D}^{\dagger}_{E}\left|\bar{\rho}\right] \tag{3.1}$$

where E is the energy of the next emission and E' the energy of the previous emission. For the first emission, $E' = Q$, which is the hard scale. The phase-space element of the emitted gluon is also included at each step. We note that for the reminder of this work we use a fixed coupling $\alpha_s = 0.118$, although the `CVolver` implementation can easily extend to a running coupling. A running coupling is formally needed to achieve full LL accuracy. This expression constitutes a map from the colour flows $\{\sigma, \bar{\sigma}\}$ to $\{\rho, \bar{\rho}\}$, and it lies at the core of `CVolver`.

The implementation we discuss first is optimized for the gaps-between-jets observable, which was discussed in Section 1.4, where emissions above a certain energy scale are vetoed within an interjet region, which we refer to as the gap. We call this implementation the 'dedicated mode" of `CVolver`. The measurement function of this non-global observable is given by

$$u_n(\{k\}_n) = \prod_{i=1}^{n} u_i(k_i) = \prod_{i=1}^{n}\left(\Theta_{\text{out}}(k_i) + \Theta_{\text{in}}(k_i)\Theta(E_i < \rho)\right), \tag{3.2}$$

where ρ is the veto scale, $\Theta_{\text{out}}(k_i)$ is unity if k_i is emitted in the out-of-gap region, and zero otherwise, and the opposite for $\Theta_{\text{in}}(k_i)$. We can safely set the evolution cutoff to the veto scale, $\mu = \rho$, since all contributions cancel below that scale, as shown in Section 2.1. Therefore, if the sampled phase-space point of an emission falls in the veto region, the event can be terminated and its weight discarded.

`CVolver` can also function in "event generator mode", in which no measurement function is applied during the evolution, and observables are measured after the event has terminated. It is simple to modify the implementation of `CVolver` to a general purpose one, which is discussed in Section 3.3.2.

Algorithm 1 describes a single event from a Monte Carlo implementation of Eq. (3.1),

in the dedicated mode. It takes as input the maximum number of emissions to consider, n_{max}, the lowest veto scale to evolve down to, ρ_c, the angular definition of the gap Ω_{gap}, a parameter λ to set the size of the collinear cutoff, the maximum order d for which to evaluate the Sudakov operator expansion in Eq. (2.31), a parameter m to set a maximum suppression in N_c, and the external legs and colour structure of the hard process. These are the most significant parameters that guide the evolution, however the full list of parameters is much longer, and is given in Section 3.3

The symbol ϕ represents the set of external legs, while the subscript indicates the number of emissions. It contains the momentum and colour representation of each parton. The functions α_0 and $\bar{\alpha}_0$ are used to sample the initial basis tensors in the amplitude and conjugate amplitude. The functions α_n and $\bar{\alpha}_n$ sample the tensors after the action of the Sudakov operator, and $\alpha_{n+1}, \bar{\alpha}_{n+1}$, the tensors after a real emission. The function E_n samples the energy scale of the next emission, while E_ρ samples the veto scale to evolve down to, between ρ_c and 1. Each sampling method has a weight W associated with it. The cumulative weight of the event is w, which includes contributions from the matrix element evaluations, as well as sampling weights required to cancel the unphysical sampling distributions.

This algorithm outlines the structure we have described thus far: we select the hard scattering density matrix to evolve from (lines 1-3), we select the veto scale of the event (lines 4-5), and then we enter a loop of emissions (line 6). For each emission, colour flows after the virtual evolution are sampled, and the matrix element of the Sudakov is computed (lines 7-12). Then, the kinematic phase-space and colour flow of the emission are selected, and its matrix element computed (lines 17-20). The loop is repeated after setting the new number of emissions and the new scale to evolve from (lines 22-24). The `while` loop ends and the weight of the event added to the cross-section if the maximum number of emissions, or the veto scale, are reached (lines 13-16).

On top of this, there are a series of other veto conditions that can cause the event to be discarded. There are three of them, and the points at which these conditions are checked are indicated with coloured horizontal lines on the algorithm. The colour and emission cutoff vetoes were implemented into `CVolver` by the author of this thesis.

The red line indicates the gap veto, which was discussed previously, and causes the event to be discarded if an emission occurs out of gap, as per the measurement function. It is evaluated after the direction of the new emission is sampled.

The blue line indicates the emission cutoff veto, which stops `CVolver` from evolv-

Algorithm 1: Evolution algorithm for a single event, dedicated mode for the gap veto cross-section.

Input : n_{max}, ρ_c, Ω_{gap} λ, d, m and hard process: $\{\phi_0, \mathbf{H}\}$
Output : ϕ, w

$\phi \leftarrow \phi_0, E' \leftarrow Q_H, n \leftarrow 0$;
$\{\sigma, \bar{\sigma}\} \leftarrow \{\alpha_0(\phi, \mathbf{H}), \bar{\alpha}_0(\phi, \mathbf{H})\}$;
$M \leftarrow [\sigma| \mathbf{H} |\bar{\sigma}], w \leftarrow W_{\text{hard}}(\{\sigma, \bar{\sigma}\}, \mathbf{H}), S \leftarrow \langle\sigma| \bar{\sigma}\rangle$;

$\rho \leftarrow E_\rho(\rho_c)$;
$w \leftarrow w \times W_{E_\rho}(\rho_c)$;
while $n \leq n_{max}$ ***and*** $u_n(\phi) > 0$ **do**
  $\{\tau, \bar{\tau}\} \leftarrow \{\alpha_n(\phi, \{\sigma\}, d), \bar{\alpha}_n(\phi, \{\bar{\sigma}\}, d)\}$;
  $w \leftarrow w \times W_{\alpha_n}(\phi, \{\sigma, \tau\}, d) \times W_{\bar{\alpha}_n}(\phi, \{\bar{\sigma}, \bar{\tau}\}, d)$;
  $E \leftarrow E_n(\phi, E', \{\tau, \bar{\tau}\})$;
  $w \leftarrow w \times W_{E_n}(\phi, E, E', \{\tau, \bar{\tau}\})$;
  $M \leftarrow [\tau| \mathbf{V}_{E,E'} |\sigma\rangle \times M \times \langle\bar{\sigma}| \mathbf{V}^\dagger_{E,E'} |\bar{\tau}]$;
  $S \leftarrow S \times \langle\tau| \bar{\tau}\rangle / S$;
  if $n = n_{max}$ ***or*** $E = \rho$ **then**
    $w \leftarrow w \times \text{Re}(MS)$;
    return $\{\phi, w\}$
  $\phi \leftarrow \Phi_{n+1}(\phi, \lambda)$;
  $w \leftarrow w \times W_{\Phi_{n+1}}(\phi, \lambda)$;
  $\{\rho, \bar{\rho}\} \leftarrow \{\alpha_{n+1}(\phi, \{\tau, \bar{\tau}\}), \bar{\alpha}_{n+1}(\phi, \{\tau, \bar{\tau}\})\}$;
  $w \leftarrow w \times W_{\alpha_{n+1}}(\phi, \{\tau, \bar{\tau}\}) \times W_{\bar{\alpha}_{n+1}}(\phi, \{\tau, \bar{\tau}\})$;
  $M \leftarrow [\rho| \mathbf{D}_E |\tau\rangle \times M \times \langle\bar{\tau}| \mathbf{D}^\dagger_E |\bar{\rho}]$;
  $S \leftarrow S \times \langle\rho| \bar{\rho}\rangle / S$;
  $E' \leftarrow E$;
  $\sigma \leftarrow \rho, \bar{\sigma} \leftarrow \bar{\rho}$;
  $n \leftarrow n + 1$
end

Vetos:

Gap veto: If $\Phi_{n+1}(\phi, \lambda) \in \Omega_{\text{gap}}$, terminate the event and set $w = 0$.

Colour veto: If $T > m$, where $1/N_c^T$ is the accumulated colour suppression, terminate the event and set $w = 0$.

Emission cutoff veto: If emission is not from a ring or string and $\Theta_{\text{rap}}(\Phi_{n+1}, \lambda) = 0$, terminate the event and set $w = 0$.

ing events where an emission from a dipole has occurred into the region removed by the collinear cutoff. This veto is discussed in Section 3.1.9.

The green line is the colour veto. At each step in the evolution the overall factor of $1/N_c^T$ is calculated. If $T > m$, where m is a maximum limit set as an input to the evolution, the evolution has steered too far away from leading colour and gets terminated. This is fundamental to the convergence of our results, and is a central part of the algorithm. It relies on being able to keep track of every factor of N_c with each matrix element insertion, and the important result that the Sudakov and real emissions operators can never bring the evolution closer to leading colour, i.e., at each step, T can only stay constant or increase. We will also keep track of every factor of N_c that we need to count as we discuss every step.

Now that we have detailed the overarching structure, we can proceed by following in detail each line in Algorithm 1. We will state explicitly how `CVolver` handles every sampling distribution and matrix element calculation.

3.1.1 Selecting the Hard Scatter Matrix element

The evolution starts by selecting the kinematics and colour flows of the hard scattering process. This corresponds to lines 1-3 in Algorithm 1.

There are currently seven different hard processes implemented in `CVolver`. For each of these, there is an automatic setting for the colour flows and different options for the kinematics. The available processes are shown in Table 3.1.

The colour flows can be set in two ways. The first is to set them manually for each run, with the `-initialFlows` command line argument. For example, a simulation with the settings `-hardProcess 4 -initialFlows 10 01` will run the $q\bar{q} \to q\bar{q}$ process, where the amplitude starts from the colour flow $|10\rangle$ and the conjugate amplitude from $|01\rangle$, as per the colour line labels in Fig. 4.9. Setting the colour flows manually as a run parameter is generally the most useful way to run `CVolver`, since each matrix element can appear with different kinematic prefactors in different Feynman diagrams, see again Fig. 4.9. After generating a set of runs for each contributing matrix element, the cross-section for each diagram is found by combining the runs from each set of colour flows with the appropriate prefactors. Manually setting the hard process colour flows also introduces no unnecessary weights due to their sampling.

The second method to select colour flows is to select the Feynman diagram, which is

Hard Process	`ProcessOption`	Number of colour lines	Number of colour flows	Number of distinct, contributing matrix elements
$Z \to q\bar{q}$	1	1	1	1 diag.
$H \to gg$	2	2	2	2 diag. + 1 interf.
$Z \to q\bar{q}q\bar{q}$	3	2	2	2 diag. + 1 interf.
$q\bar{q} \to q\bar{q}$	4	2	2	2 diag. + 1 interf.
$q\bar{q} \to l\bar{l}$	5	1	1	1 diag.
$gg \to gg$	6	4	24	6 diag. + 15 interf.
$qg \to qg$	7	3	6	5 diag. + 10 interf.

Table 3.1: Hard processes currently implemented in `CVolver`. The `ProcessOption` parameter is the name given to each of the processes within the code. The contributing matrix elements for the $2 \to 2$ processes consider all gluon exchange channels. We do not double count interferences, since the evolution is symmetric between amplitude and conjugate amplitude. For the colour flow contributions to each channel in $qg \to qg$ and $gg \to gg$ see Figs. 4.10a and 4.10b.

implemented for each process. These sample over the contributing colour matrix elements, with a weight according to their prefactor in the full hard process. We can consider as an example $H \to gg$. There are $n_g! = 2!$ possible colour flow configurations, and so the density matrix has four contributing terms:

$$\mathbf{H}_{gg} = |10\rangle\langle 10| - \frac{1}{N_c}|01\rangle\langle 10| - \frac{1}{N_c}|10\rangle\langle 01| + \frac{1}{N_c^2}|01\rangle\langle 01|, \tag{3.3}$$

where the states with singlet gluons have a factor of $-1/N_c$ associated with them as per the Fierz identity. The two interference terms are equivalent, and we do not need to evolve both of them separately. The initial colour flows are selected following the discrete probability distribution:

$$p_{gg}(\sigma, \bar{\sigma}) = \frac{|[\sigma|\,\mathbf{H}_{gg}\,|\bar{\sigma}]|\,N_c^2}{(N_c+1)^2}. \tag{3.4}$$

and afterwards the weight, hard scattering matrix element, and initial scalar product matrix are assigned:

$$W_{\text{hard}} = \frac{1}{p_{gg}(\sigma,\bar{\sigma})}, \qquad M_{\sigma\bar{\sigma}} = [\sigma|\,\mathbf{H}_{gg}\,|\bar{\sigma}], \qquad S_{\sigma\bar{\sigma}} = \langle\sigma|\,\bar{\sigma}\rangle. \tag{3.5}$$

An equivalent distribution is implemented for every other process. In the case of $2 \to 2$ scattering, the preset diagram is always t-channel gluon exchange. Other diagrams are simple to implement, but it has not been done yet. Setting the hard process colour flows manually is generally more computationally efficient, as sometimes the same colour matrix element contributes to different Feynman diagrams with different colour prefactors.

The evaluation of $S_{\sigma\bar{\sigma}}$ and $M_{\sigma\bar{\sigma}}$ constitute the first source of colour suppression $1/N_{\mathrm{c}}^{p}$. The colour veto checks p against the colour suppression limit m and terminates the evolution if $p > m$. Normally we want to include all interference contributions to each diagram, thus m is usually chosen so that the colour veto always passes at this stage. Also note that if the colour flows were set manually, then $M_{\sigma\bar{\sigma}} = W_{\text{hard}} = 1$, and p can only be non-zero if the colour flows correspond to an interference.

The kinematics depend on the choice of process. For the two jet processes, their normalised four-momenta are set to

$$n_a = (1, 0, 0, 1) \qquad n_b = (1, \sin\theta, 0, \cos\theta), \tag{3.6}$$

where θ is set with the `-dipoleangle` command line argument, and the preset is for the jets to be back-to-back.

The $2 \to 2$ kinematics can be set in two ways. In both cases the initial-state jets are back-to-back. Using the `-dipoleangle` argument, the jets are defined with the four-momenta:

$$\begin{aligned} n_a &= (1, 0, 0, 1) & n_b &= (1, 0, 0, -1) \\ n_c &= (1, \sin\theta, 0, \cos\theta) & n_d &= (1, -\sin\theta, 0, -\cos\theta) \end{aligned} \tag{3.7}$$

where a, b are incoming, c, d are outgoing, and again `-dipoleangle` sets θ. Alternatively, the final state can be defined to not be back-to-back:

$$\begin{aligned} n_a &= (1, 0, 0, 1) & n_b &= (1, 0, 0, -1) \\ n_c &= (1, \operatorname{sech} y_c, 0, \tanh y_c) & n_d &= (1, \operatorname{sech} y_d, 0, \tanh y_d) \end{aligned} \tag{3.8}$$

where y_c, y_d are set through the `-jet-rapidities` command line argument. These two settings correspond to the two kinematic configurations we use for the results in Chapter 4.

The definition of the gap region depends on how the kinematics were defined. For the $0 \to 2$ processes, or the $2 \to 2$ processes when using the `-dipoleangle` argument, the gap is defined as the region $-\cos\theta_g < \cos\theta < \cos\theta_g$, where the angles are with respect to the beam axis $(1, 0, 0, 1)$, and θ_g is set with the `-cone` command line argument. Figures for the $0 \to 2$ and $2 \to 2$ cases are shown in Figs. 4.1 and 4.8 respectively.

If the kinematics were set with `-jet-rapidities`, then the gap region is defined as $y_d + R < y < y_c - R$, where R is set with the `-cone` argument. This command line argument is important to obtain results efficiently for a given gap size, as evolution is terminated when emitting into the gap region (the red veto between lines 16 and 17). Since the measurement function for small gap would set the weight of a subset of events to zero, the `Analyses` part of the implementation, where all observables and histograms are defined, can generate any cross-section from a sample with a larger gap region than the one defined with `-cone`. There is also a special case: setting `-cone -1` removes the gap altogether, accepting all emissions. We have used this setup extensively to test the unitarity of the algorithm.

These are the preset configurations and the already-implemented command line arguments, however it is quite simple to code directly into `HardProcessMatrix` any kinematics or other particles. The colour algebra and the evolution are completely modularised, and can handle any hard process. The options we have described were implemented into `CVolver` by the author of this work to facilitate the studies in Chapters 4 and 5.

3.1.2 Setting the evolution cutoff

In `CVolver`, the energy of every particle in the process is normalised with respect to the hard scale Q_H. Thus, the initial scale to start evolution from Q_H is defined as 1, with subsequent emissions having energies between 0 and 1.

In dedicated mode a cutoff scale for the evolution must be chosen for each event. This will be the energy cutoff below which the observable is inclusive of all emissions, ρ. Different evolution cutoffs must be chosen for each event in order to populate the observable $\Sigma(\rho)$ for different ρ, down to ρ_c which is set as an external parameter of the run. This method corresponds to lines 4 and 5 in Algorithm 1.

The evolution cutoff is sampled between ρ_c and 1 with a probability density function

$$p(\rho) = \begin{cases} -\dfrac{1}{\ln \rho_c}\rho^{-1} & \alpha = 1 \\ \dfrac{-\alpha+1}{1-\rho_c^{-\alpha+1}}\rho^{-\alpha} & \alpha > 1 \end{cases} \tag{3.9}$$

where $\alpha \geq 1$ is a parameter that can be increased to prioritise sampling events with low evolution cutoffs. In general, lower values of ρ require higher statistics because higher multiplicities contribute more, and these require exploring a larger number of trajectories in colour space. The cutoff is sampled by inversion, and uses a source of random numbers drawn from a flat distribution within the range $[0, 1]$, denoted $\mathcal{R}$. Thus, the cutoff is selected from

$$\rho = E_\rho(\rho_c) = \begin{cases} \rho_c^{1-\mathcal{R}} & \alpha = 1 \\ \left(\mathcal{R} + (1-\mathcal{R})\rho_c^{-\alpha+1}\right)^{1/(-\alpha+1)} & \alpha > 1. \end{cases} \tag{3.10}$$

This process also contributes a weight to the event:

$$W_{E_\rho} = \frac{1}{p(\rho)} = \begin{cases} \rho \ln \dfrac{1}{\rho_c} & \alpha = 1 \\ \rho^\alpha \dfrac{1-\rho_c^{-\alpha+1}}{-\alpha+1} & \alpha > 1. \end{cases} \tag{3.11}$$

The parameter α constitutes our first lever to manipulate the sampling distributions within `CVolver`. These sampling parameters are fundamental for the Monte Carlo to converge, since the available phase-space is computationally astronomical, and it is necessary to steer the evolution towards the regions with significant contributions. They were implemented in order to improve convergence for the results presented in Chapters 4 and 5. As we continue detailing each step, we will find these parameters appear in many sampling distributions. We will discuss their effects, and sensible choices for them in Section 3.3.1.

With the hard scattering matrix element calculated, the initial colour flows chosen, and the evolution cutoff determined, the next step in the algorithm is to enter the evolution loop, which starts on line 6 of Algorithm 1.

3.1.3 Selecting flows after the virtual evolution

The virtual evolution consists of three steps, as shown in Algorithm 1: the new colour flows in the amplitude and conjugate amplitude after the action of the Sudakov operator must be selected (lines 7 and 8). Then, the energy of the next emission is sampled, which specifies the scale to evolve down to (lines 9 and 10). Finally, the matrix element contributions are calculated and multiplied by the amplitude and conjugate amplitude (lines 11 and 12).

This section follows the explanation given in [24], although some modifications to the algorithm have been made, and the notation has been changed to closer reflect what is currently implemented in `CVolver`.

The $\mathrm{N}^d\mathrm{LC}'$ approximation of the Sudakov operator in Eq. (2.31) has the advantage that it limits the list of colour flow vectors accessible after the action of the Sudakov operator to those at most d swaps away from the initial flow. Therefore, we only have to sample from the colour flows that satisfy this condition, significantly reducing the number of possible trajectories, and greatly simplifying the computational complexity. The parameter d is set through the command line as is one of the main input parameters to Algorithm 1.

The number of swaps d' is sampled according to an exponentially falling distribution. Therefore, elements closer to the colour diagonal, which are not colour suppressed, are prioritised. After this, the level-swap algorithm, detailed later, is used to select a permutation which is d' swaps away. These steps are performed independently in the amplitude and conjugate amplitude.

Before proceeding with the sampling, all singlet flows are taken out of the colour flow permutations in the amplitude and conjugate amplitude, reducing the number of colour lines from $n_\sigma = n_{\bar{\sigma}}$ to n'_σ and $n'_{\bar{\sigma}}$, which may not be equal. Singlets do not carry colour nor anti-colour, and therefore cannot connect to any virtual gluon exchanges or real emissions. They do not participate in the subsequent evolution, and can be referred to as sterile modes.

To determine the number of swaps d', `CVolver` uses sampling by inversion from the continuous probability distribution

$$p(x) = \frac{\ln N_c^s}{1 - N_c^{-s(d_{\text{max}}+1)}} N_c^{-sx}, \tag{3.12}$$

for x between $d_{\text{min}} = 0$ and $d_{\text{max}} = \min(d, n'_\sigma - 1) + 1$ for the amplitude, or $d_{\text{max}} = \min(d, n'_{\bar{\sigma}} - 1) + 1$ for the conjugate amplitude. We have also introduced the second

sampling parameter s, which is set through the command line argument `-dSup`. Its purpose is to suppress or enhance the number of swaps during evolution, and in the former case keep the evolution from straying too far off the diagonal. It is particularly important for the $2 \to 2$ processes, and we discuss it further in Section 3.3.1.

The definition of $d_{\max}$ accounts for the fact that for the first steps of the evolution, it is possible that the maximum number of possible swaps $n'_\sigma - 1$ or $n'_{\bar{\sigma}} - 1$ is smaller than the externally set parameter d. The additional unit is necessary since we take d' to be the floor of the continuous variable: $d' = \lfloor x \rfloor$. Thus the probability of sampling d' is given by

$$p(d') = \int_{d'}^{d'+1} p(x)\mathrm{d}x = \frac{N_c^{sd_{\max}}(1 - N_c^s)}{1 - N_c^{-s(d_{\max}+1)}} N_c^{-sd'}, \tag{3.13}$$

with an associated weight $W_d = 1/p(d')$, which will be included within W_{α_n} and $W_{\bar{\alpha}_n}$.

Once the number of swaps d' has been chosen, it is necessary to find all vector flows which are d' swaps away from the colour state the Sudakov operator will be acting on. This is done with the level swap algorithm, represented in Figure 3.2.

We can walk through each step on the diagram to explain the method to generate the accessible basis tensors. The 0^{th} layer corresponds to the initial colour permutation, $|\sigma\rangle$. The first layer consists of a subset of intermediate tensors which are one swap away from $|\sigma\rangle$, and from which all the tensors two swaps away can be accessed by a particular swap. Each layer of swaps is achieved by first picking an index i, denoted by the red numbers above each branch, as the colour index corresponding to one of the pairs of anti colour indices we wish to swap. In general, i runs from 0 to $n'_\sigma - 1$, but at each given layer the only values allowed are from $d' - l - 1$ to $i_l - 1$, where l denotes the layer of branching, and i_{l-1} is the i value of the branch in the layer above the one being considered, with $i_0 = n'_\sigma - 1$. The second anti-colour we wish to swap with, $i + k$, can take any value above i, so k is in the range $[1, n'_\sigma - i - 1]$. The number of possible values of k is indicated by the blue number in each bubble. The layering process is repeated, where each i corresponds to a new branch, until layer d' is reached.

By applying to $|\sigma\rangle$ all the swaps described by following each of the branches (where each swap of an intermediate layer is combined with the following swaps), the whole set of permutations precisely d' swaps away from $|\sigma\rangle$ is generated. Therefore, the total number of such permutations can be found by summing all the possible k indices on the last layer, multiplied by the number of k indices on each previous branch. It is necessary to determine the multiplicity of permutations in each branch in order to sample tensors in an unbiased

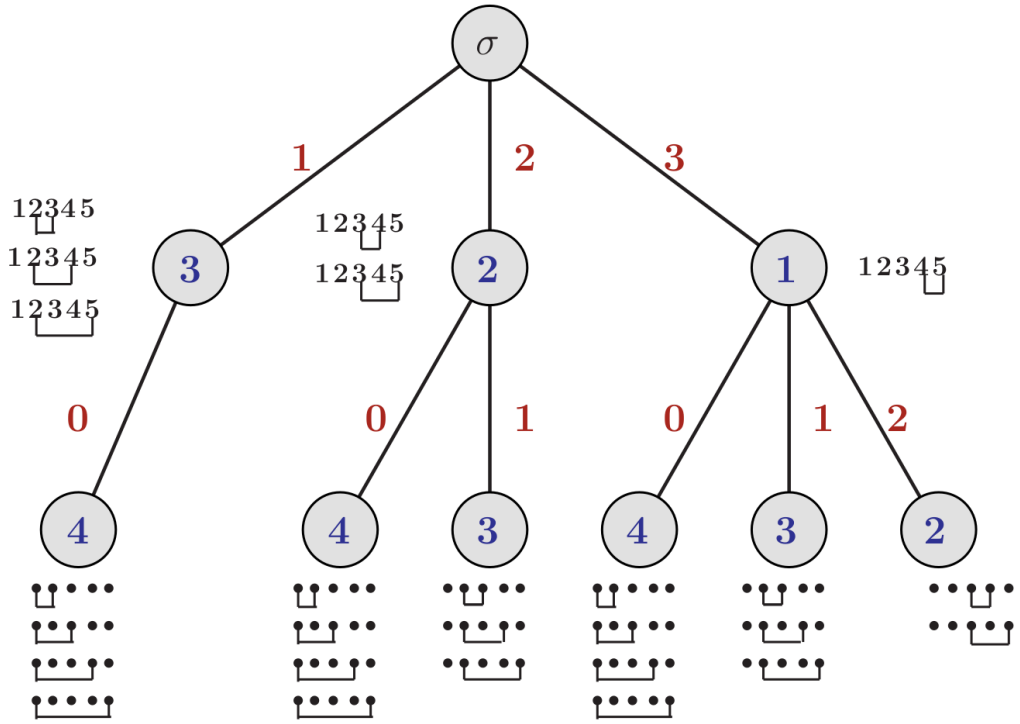


Figure 3.2: Pictorial representation of the level swap algorithm, which samples all tensors that are exactly two swaps away from $|\sigma\rangle = |12345\rangle$. The red number above each branch indicates the first of the two indices being swapped, called i_l. The total number of indices k, which represent the second index partaking in the swap, is written in blue. For each layer and branch the possible swaps that occur are shown explicitly, on the first layer using the specific anti-colour indices, and on the second layer with dots. Figure taken from [24].

manner. This is given by the expression

$$M_{n,d'}(l, i_l) = \sum_{j=d'-l-1}^{i_l-1} (n-1-j) \times M_{n,d'}(l+1, j), \tag{3.14}$$

where $M_{n,d'}(l, i_l)$ is the multiplicity on layer l and branch i_l for the process with n colour lines (excluding singlets) and d' swaps. This expression can be used recursively until $M_{n,d'}(l = d', j) = 1$ is reached. To illustrate how this formula works, let us use it to calculate the total number of tensors in the final layer of Figure 3.2. This is given by

$$\begin{aligned} M_{5,2}(l=0, i_0=4) &= 3 \times M_{5,2}(1,1) + 2 \times M_{5,2}(1,2) + 1 \times M_{5,2}(1,3) \\ &= [3 \times 4] + [2 \times (4+3)] + [1 \times (4+3+2)] = 35, \end{aligned} \tag{3.15}$$

which gives the same result as counting directly the number of tensors along each branch

in Figure 3.2.

Having found the list of accessible tensors and a way to count them, it is possible to select the next colour flow in the evolution. This is done recursively by following a single branch at each layer of the level swap algorithm: starting from the 0^{th} layer, a branch i is selected with probability

$$P_{n,d',l}(i) = \frac{(n-i-1)M_{n,d'}(l+1,i)}{M_{n,d'}(l,i_l)} \tag{3.16}$$

which gives each i a probability proportional to the number of tensors along that branch. After an i is selected, k is chosen from a flat distribution in the range of possible values, from 1 to $n-i-1$. The anti-colour indices associated with colour indices i and k are then swapped. This process is repeated, going layer by layer, until there have been d' swaps. The final step is to reintroduce in the tensor the singlet flows that were taken out at the beginning.

Contextualising this within Algorithm 1, this method corresponds to the functions $\alpha_n(\sigma)$ and $\bar{\alpha}_n(\bar{\sigma})$, which select the colour flows τ and $\bar{\tau}$ in the amplitude and conjugate amplitude respectively after virtual evolution. The two functions are equal, but need to be performed independently as the initial colour states σ and $\bar{\sigma}$ are in general not the same.

Since all possible flows are considered with equal probability, the weight associated with this sampling is given by the total number of accessible tensors. To obtain the full weight associated with the selection of colour flows after virtual evolution, we need to combine the weights from sampling the number of swaps and the new basis vector:

$$W_{\alpha_n} = W_d \frac{1}{p(\tau)} = \frac{1-N_c^{-s(d_{\max}+1)}}{N_c^{sd_{\max}}(1-N_c^s)} N_c^{sd'} \times M_{n'_\sigma,d'}(l=0, i_0 = n'_\sigma - 1) \tag{3.17}$$

and equivalently for the conjugate amplitude.

3.1.4 Sampling the scale of the next emission

Having selected the colour flows after evolution, the next step in Algorithm 1 is to choose what scale to evolve down to, lines 9 and 10. This selection has a deep impact on the shower evolution, as it also dictates the energy of the next emission, and thus it controls the multiplicity of the sampled events at each energy scale.

At each step in the evolution the best choice of sampling distributions are those that follow the physical distributions themselves. For this reason, we will take the real emission scale distribution directly from Eq. (3.1):

$$\mathrm{d}R(E,n,\tau,\bar{\tau}) = \frac{\alpha_s}{\pi}\frac{dE}{E}\sum_{i\neq j}\left[\tau\right|\mathbf{T}_i\left|\sigma\right\rangle\left\langle\bar{\sigma}\right|\mathbf{T}_j^\dagger\left|\bar{\tau}\right]\omega_{ij}(n)\frac{\mathrm{d}\Omega}{4\pi}, \tag{3.18}$$

which is simply the cross-section level contraction of emission operators in Eq. (2.24), with the phase-space element in Eq. (2.3). The minus sign in the latter was absorbed in the contraction of polarisation vectors to form the dipole function ω_{ij}. This distribution includes information about the energy scale, direction, and colour flows of the emission, and we want to define a probability density function to sample each of these quantities. It is possible to factorise them in order to obtain independent and normalised distributions:

$$\begin{aligned}\mathrm{d}R(E,n,\tau,\bar{\tau}) = \frac{\alpha_s}{\pi}\frac{dE}{E}\sum_{i\neq j}\xi_{ij}\widetilde{\Omega}_{ij}\times\left(\frac{\omega_{ij}(n)}{\widetilde{\Omega}_{ij}}\frac{\mathrm{d}\Omega}{4\pi}\right)\\ \times\left(\frac{\left[\tau\right|\mathbf{T}_i\left|\sigma\right\rangle\left\langle\bar{\sigma}\right|\mathbf{T}_j^\dagger\left|\bar{\tau}\right]}{\left|\left[\tau\right|\mathbf{T}_i\left|\sigma\right\rangle\left\langle\bar{\sigma}\right|\mathbf{T}_j^\dagger\left|\bar{\tau}\right]\right|}\right)\times\left(\frac{\left\langle\tau\right|\left.\bar{\tau}\right\rangle}{\left\langle\sigma\right|\left.\bar{\sigma}\right\rangle}\frac{\left|\left[\tau\right|\mathbf{T}_i\left|\sigma\right\rangle\left\langle\bar{\sigma}\right|\mathbf{T}_j^\dagger\left|\bar{\tau}\right]\right|}{\xi_{ij}}\right),\end{aligned} \tag{3.19}$$

where $\widetilde{\Omega}_{ij}$ is the quantity we use to normalise the radiation pattern distribution, and will be defined in Sec. 3.1.7. The quantity ξ_{ij} is given by:

$$\xi_{ij} = \sum_{\tau',\bar{\tau}'}\frac{\left\langle\tau'\right|\left.\bar{\tau}'\right\rangle}{\left\langle\sigma\right|\left.\bar{\sigma}\right\rangle}\left|\left[\tau'\right|\mathbf{T}_i\left|\sigma\right\rangle\left\langle\bar{\sigma}\right|\mathbf{T}_j^\dagger\left|\bar{\tau}'\right]\right|, \tag{3.20}$$

which is the colour factor associated with a cross-section contracted emission, as we defined in Eq. (2.41), summed over all possible colour flows after the emission. There is a lot to unpack in Eq. (3.19). The first parenthesised quantity consists of a normalised distribution over the direction of the emission. Similarly, the third term in parentheses encodes the relative importance of the possible flows after emission, $\tau, \bar{\tau}$. It includes the change in scalar product matrix, which was not present in Eq. (3.18). Sampling distributions need to be positive definite, so any change of sign arising from the choice of colour flows is contained in the second parenthesis. These are the distributions that will be used to sample the momentum and colour flows of the next emission.

Discussing these ingredients would be skipping a few steps ahead, since we have yet to evaluate the virtual evolution. The remaining part of Eq. (3.19), outside any parenthesis, consists of a probability distribution for emitting with energy E_k. However, this distribution is incomplete; it does not reflect the fact that no particle has been emitted between the current energy scale and the next one. To take the non-emission probability into account, we use the Sudakov veto algorithm, which samples according to the distribution:

$$\mathrm{d}S_P(E|E') = P(E)\exp\left(-\int_E^{E'} P(q)\mathrm{d}q\right)\mathrm{d}E = P(E)\Delta_P(E|E')\mathrm{d}E, \tag{3.21}$$

where E' is the scale of the previous emission, or $E' = Q_H = 1$ for the first emission. The first factor, $P(E)$, constitutes the probability for emitting a soft gluon with energy E, while $\Delta_P(E|E')$ is the probability for not emitting any particles between energy scales E and E'. We can identify $P(E)$ with the energy distribution found in $\mathrm{d}R(E_k, n, \tau, \bar{\tau})$, and we will add a few more parameters to steer the sampling:

$$P(E) = \frac{\alpha_s}{\pi}\frac{\left(W_V\Theta_{W_V} + (1-\Theta_{W_V})\right)}{\left(\beta\Theta_\beta(n) + (1-\Theta_\beta(n))\right)}\frac{1}{E}\sum_{i\neq j}\xi_{ij}\widetilde{\Omega}_{ij}. \tag{3.22}$$

where the second fraction includes the introduced sampling parameters, and the rest was taken directly from $\mathrm{d}R(E_k, n, \tau, \bar{\tau})$. The β parameter is our third lever, it can be set through the command line argument `-beta`, and its purpose is to manipulate the probability of selecting higher or lower scales, and therefore enhancing or suppressing the rate of emissions. It is accompanied by $\Theta_\beta(n)$, which is a switch that can activate the β parameter at some multiplicity n, which is set through the command line argument `-beta-switch`. These parameters are necessary in order to enhance the statistics of certain multiplicities at certain energy scales.

On the numerator we have W_V, which is the weight associated with the selection of colour flows after the Sudakov operator: $W_V = W_{\alpha_n} \times W_{\bar{\alpha}_n}$. We can also absorb this factor from the virtuals into the scale selection as a way to reduce the event weights. Whether to do this or not is set through the `-virtual-weights-in-proposal` command line argument, which turns Θ_{W_V} on or off. This is our fourth lever for steering the evolution and will be discussed, along the β parameter, in Section 3.3.1.

One might notice that this distribution has not been taken faithfully from Eq. (3.19); $\sum_{i\neq j}\xi_{ij}\widetilde{\Omega}_{ij}$ cannot be factored out since it is followed by other ij-dependent terms, in the

second and third parentheses. Fundamentally, the problem arises from the cross-section contracted real emission distribution involving a sum over all dipole pairs $i \neq j$, when in our Monte Carlo algorithm we are choosing a single trajectory in colour and kinematics. We can solve both problems at once by selecting an emitting dipole from the distribution:

$$P_{\text{dip}}(i,j) = \frac{\xi_{ij}\widetilde{\Omega}_{ij}}{\sum_{i\neq j}\xi_{ij}\widetilde{\Omega}_{ij}}, \tag{3.23}$$

where the denominator cancels part of the weight introduced from selecting the energy scale, and the numerator will cancel with the normalisations of the direction and colour flow distributions respectively.

Returning to the energy scale selection, we use sampling by inversion from $\mathrm{d}S_P(E|E')$, Eq. (3.21), using a source of equal probability random numbers $\mathcal{R} \in (0,1)$, to select the energy scale of the next emission:

$$E = E'(1-\mathcal{R})^{\left(\beta\Theta_\beta(n)+(1-\Theta_\beta(n))\right)\pi/\left[\alpha_s\left(W_V\Theta_{W_V}+(1-\Theta_{W_V})\right)\sum_{i\neq j}\xi_{ij}\widetilde{\Omega}_{ij}\right]}. \tag{3.24}$$

However, the scale sampled here cannot be immediately accepted, and the weight of the selection depends on how the evolution proceeds. Section 3.1.5 will focus on this, and only after validating the scale can we proceed to evaluate the Sudakov matrix elements (lines 10-12 in Algorithm 1).

We have not described yet the calculation of ξ_{ij} and $\widetilde{\Omega}_{ij}$ for each possible pair ij. The former is discussed in Section 3.1.8, since it follows the same calculation used for sampling the after-emission colour flows. The latter is dependent on the collinear cutoff prescription, which we have alluded to a few times, but not defined properly so far. Clearly the integral over ω_{ij} diverges as the emission becomes collinear with i or j, although we know these infrared poles will cancel against the equivalent poles in the phase-space integral of the real emissions. We need to cut them out of each individual piece in order to obtain convergent matrix elements within individual Monte Carlo events. The collinear cutoff prescription is defined in Section 3.1.6. For now, it suffices to note that $\widetilde{\Omega}_{ij}$ is finite, so we can proceed with evaluating the Sudakov matrix element.

3.1.5 Validating the emission scale and evaluating the evolution operator element

In reality, the expression given for $\mathrm{d}S_P(E|E')$ in Eq. (3.21) is incomplete; it does not take into account the maximum number of emissions $n_{\max}$, nor the cutoff scale ρ, below which particles should not be emitted. A complete and normalised distribution is given by:

$$\begin{aligned}\mathrm{d}S_P(\rho, E|E') = &\Delta_P(\rho|E')\delta(E-\rho)\mathrm{d}E \\ &+ (1-\Delta_P(\rho|E'))\,\delta_{n,n_{\max}}\delta(E-\rho)\mathrm{d}E \\ &+ P(E)\Delta_P(E|E')(1-\delta_{n,n_{\max}})\Theta(E'-E)\Theta(E-\rho)\mathrm{d}E.\end{aligned} \tag{3.25}$$

The details of its construction are found in [106]. For our purposes it is enough to understand where each line comes from. The third line is inherited from Eq. (3.21): for energy scales $\rho < E < E'$, the probability of the next emission occurring at energy E is the product of the non-emission probability between E' and E, $\Delta_P(E|E')$, and the probability of emitting with energy E, $P(E)$. Notice that this term applies to all emissions, except when the maximum number of emissions has been reached, and only within the energy range set by the step functions.

If the maximum multiplicity has not been reached, but the sampled E is below ρ, only the first line of the distribution survives, which tells us to discard the selected scale and set $E = \rho$, with an associated probability for not emitting down to the evolution cutoff, $\Delta_P(\rho|E')$.

On the other hand, if the maximum multiplicity has been reached, i.e., $n = n_{\max}$, the first and second lines partly cancel and we are forced to select $E = \rho$ with probability 1.

We can now go back to the scale E selected naively from Eq. (3.21) and correct it according to Eq. (3.25). There are four cases:

Case 1: $E > \rho, n \neq n_{\max}$
The scale is accepted, and the sampling process is equivalent to multiplying the weight by $P(E)\Delta_P(E|E')$. We keep the factor related to the emission operators, but we divide out the non-emission probability $\Delta_P(E|E')$ from the weight since we want to use instead the correct Sudakov matrix elements. We also have to divide out the unphysical emission rate parameter and the weight from virtual flows selection, $\big(W_V\Theta_{W_V}+(1-\Theta_{W_V})\big)/\big(\beta\Theta_\beta(n)+(1-\Theta_\beta(n))\big)$. If $\Theta_{W_V} = 1$, then W_V gets divided out from the weight at this stage and

it will cancel the weight applied on line 8 of Algorithm 1, thus completely absorbing the virtual flows selection weight into the scale proposal, and reducing the overall weight of the event.

Case 2: $E < \rho, n \neq n_{\text{max}}$
This emission is forbidden, so we set $E = \rho$ and virtually evolve the event to the cutoff scale. This selection is made with probability $\Delta_P(\rho|E')$, and we divide it out of the weight in order to use instead the correct Sudakov matrix element.

The last two cases are more subtle.

Case 3: $E < \rho, n = n_{\text{max}}$
At this stage in the evolution, the event can only be virtually evolved to the cutoff, since no more emissions are allowed. Thus, the selected scale is not entirely relevant, and only serves to indicate whether the event would actually terminate at this multiplicity, if $E < \rho$ is sampled from the distribution, or if the event is being forcefully made not to emit. As such, we can associate the weight of these events with any constant $0 < a < 1$, which is simply done by multiplying the weight by $a/\Delta_P(\rho|E')$.

Case 4: $E > \rho, n = n_{\text{max}}$
We require that events in this case carry a total weight $1 - a$, in order to maintain normalisation. However, cancelling the selection weight is more complicated in this case, since it is given by the integral over $P(E)\Delta_P(E|E')$. To do so, we modify the weight of each event with the terms in red:

$$\int_{\rho}^{E'} P(q)\Delta_P(q|E')\mathrm{d}q \times \frac{1}{\Delta_P(q|E')P(q)} \frac{\mathrm{d}f(q)}{\mathrm{d}q} \frac{1-a}{f(E') - f(\rho)} = 1 - a, \tag{3.26}$$

where $f(q)$ can be any continuous real-valued function defined for $q \in [\rho, E']$. After running the Monte Carlo and sampling over the region $\rho < E < E'$, the total probability associated with this scenario integrates to $1 - a$.

The choice of a and $f(q)$ is arbitrary, and they only serve to weight differently events which have the same behaviour: evolving down to ρ and terminating. In previous versions of `CVolver` the physical values $a = \Delta_P(\rho|E')$ and $f(q) = \Delta_P(q|E')$ were used, however

these could lead to weight fluctuations depending on the sampled scale E'. The current version uses $a = 0.5$, and $f(q) = q$. These give every possible scale selection the same weight, which improves numerical convergence. It is important to note that the selection weights are not required to be physical, and that the choice to terminate evolution and evolve down to the cutoff at a certain multiplicity is not physical either. Lastly, these weights do not usually have a large impact, as they only affect the last emission and n_{max} is usually chosen to be larger than the number of necessary emissions for the cross-section to converge at some ρ. We discuss these parameters in Section 3.3.1.

This validation process for the scale selection is contained within the function $E_n(\phi, E', \{\tau, \bar{\tau}\})$ in line 9 of Algorithm 1. The following steps indicate the evaluation of the Sudakov matrix elements and the updated scalar product matrix (lines 11 and 12). With both the colour flows and scale after virtual evolution chosen, it is relatively simple to evaluate the matrix element from the $1/N_c$-series expansion of the Sudakov operator in Eq. (2.35). As discussed previously, the Γ, ρ, and Σ components are calculated by tracking the coefficients in front of the different colour product operators in Figs. 2.3 and 2.4. An expression for Σ was given as an example in Eq. (2.38). Then, all possible sequences of colour flows $\{\sigma_0, ..., \sigma_l\}$, of length d', where $\sigma = \sigma_0$ and $\sigma_l = \tau$ must be found. Lastly, the R functions must be computed. The closed form expression for an R function for any number of flows is implemented in `CVolver`, however, in practice all cases where one or more flows become degenerate must be added in manually to avoid numerical instability. At this time, all degenerate R functions needed up to $d' = 4$ are implemented. For details on the implementation for both the closed form R functions and colour flow sequences algorithm see [24].

The evaluation of the Sudakov matrix elements and the update to the scalar product matrix brings us to the second colour veto. As a reminder, assuming no emissions have occurred yet, we have arrived at this point with a total colour suppression factor $T = p$, where T is from the accumulated colour $1/N_{\text{c}}^T$, and p is the suppression arising from the hard scatter and initial scalar product matrix elements. There are now two additional contributions to colour suppression: explicit factors of $1/N_{\text{c}}$ arising from the Sudakov operator series expansion, and further suppression from the scalar product. The former will add $s = d' + \overline{d}'$ powers of $1/N_{\text{c}}$, where d' and $\overline{d}'$ are the numbers of swaps taken in the amplitude and conjugate amplitude respectively, since each swap is associated with a factor of $1/N_{\text{c}}$. The latter source is the update to the scalar product $\langle\tau|\bar{\tau}\rangle/\langle\sigma|\bar{\sigma}\rangle = 1/N_{\text{c}}^t$.

An important result is that the combined action of the Sudakov and scalar product can never bring the evolution closer to leading colour: $s + t \geq 0$. For example, if one swap

changes the density element from $|10\rangle\langle 01|$ to $|01\rangle\langle 01|$, then the $1/N_c$ from the swap cancels the N_c from the scalar product, and $s + t = 0$.

In total, the event gets terminated if $T = p + (s + t) > m$, since the event is deemed too suppressed and not worth the computational time. Otherwise, we can continue our journey down Algorithm 1, and we find a branching point in line 13. The first option is to continue emitting, and this occurs when the energy scale sampled falls within Case 1. In Cases 2, 3 and 4, the process is evolved down to the cutoff scale and the event terminates (lines 19, 20 and 21). We will continue the path following from Case 1, and the next step is to sample the direction of the next emission, in lines 14 and 15.

3.1.6 Collinear cutoff prescription

We must first tie some loose ends and finally define the collinear cutoff prescription. This is the procedure by which we remove small cones around the direction of each jet, removing collinear singularities. We can safely do this for observables that are inclusive over all hard-collinear physics, such as the gap veto cross-section. Furthermore, if the cone size is determined by a parameter λ, then testing if our results are independent of λ for small enough values provides another cross-check.

There are many ways to implement the collinear cutoff prescription. The current definition implemented is called the rapidity-type cutoff, and it is the one used for all the results presented in Chapters 4 and 5. This cutoff is defined with the Heaviside theta function

$$\Theta_{\text{rap}} = \Theta\left(\min\left(\frac{n\cdot n_i}{n\cdot n_j}, \frac{n\cdot n_j}{n\cdot n_i}\right) - \frac{\lambda}{n_i\cdot n_j}\right), \tag{3.27}$$

where λ determines the size of the removed regions, and n is the normalised four-momentum of an emission from parents i and j. We must remove contributions that fail this condition in both the virtual angular integral, and the phase-space of the real emissions, to ensure they cancel. The quantity $\omega_{ij}\Theta_{\text{rap}}$ is difficult to integrate in the lab frame, but we can define a boost to the zero momentum frame (ZMF) of the ij pair:

$$\begin{aligned}\hat{n}_i^\mu &= \sqrt{\frac{n_i\cdot n_j}{2}}(1,0,0,1) = \left(\Lambda^{-1}(n_i\cdot n_j)\right)^\mu_\nu n_i^\nu,\\ \hat{n}_j^\mu &= \sqrt{\frac{n_i\cdot n_j}{2}}(1,0,0,-1) = \left(\Lambda^{-1}(n_i\cdot n_j)\right)^\mu_\nu n_j^\nu.\end{aligned} \tag{3.28}$$

where the vectors in the lab frame are oriented such that

$$\begin{aligned} n_i^\mu &= (1, 0, 0, 1) \\ n_j^\mu &= (1, \sin\theta_{ij}, 0, \cos\theta_{ij}). \end{aligned} \tag{3.29}$$

The virtual angular integral transforms as

$$\begin{aligned} \Omega_{ij} &= \int \frac{\mathrm{d}\Omega}{4\pi} \frac{n_i \cdot n_j}{n \cdot n_i \;\; n \cdot n_j} \Theta_{\text{rap}}\left(\frac{n \cdot n_i}{n \cdot n_j}, n_i \cdot n_j\right) \\ &= \int \frac{\mathrm{d}\hat{\Omega}}{4\pi \left[(\Lambda\hat{n})^0\right]^2} \frac{\left[(\Lambda\hat{n})^0\right]^2 \hat{n}_i \cdot \hat{n}_j}{\hat{n} \cdot \hat{n}_i \;\; \hat{n} \cdot \hat{n}_j} \Theta_{\text{rap}}\left(\frac{\hat{n} \cdot \hat{n}_i}{\hat{n} \cdot \hat{n}_j}, \hat{n}_i \cdot \hat{n}_j\right) \\ &= \int \frac{\mathrm{d}\hat{\Omega}}{4\pi} \frac{\hat{n}_i \cdot \hat{n}_j}{\hat{n} \cdot \hat{n}_i \;\; \hat{n} \cdot \hat{n}_j} \Theta_{\text{rap}}\left(\frac{\hat{n} \cdot \hat{n}_i}{\hat{n} \cdot \hat{n}_j}, \hat{n}_i \cdot \hat{n}_j\right) = \hat{\Omega}_{ij}, \end{aligned} \tag{3.30}$$

where $\hat{n} = (1, \sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$ is the normalised four-momentum of the emission in the ij-ZMF. We have shown one of the advantages of the rapidity-type cutoff: it transforms nicely into the ij-ZMF, and we can calculate the exact angular integral analytically:

$$\Omega_{ij} = \int \frac{\mathrm{d}\Omega}{4\pi} \omega_{ij} \Theta_{\text{rap}} = \ln \frac{n_i \cdot n_j}{\lambda}. \tag{3.31}$$

Therefore, applying the collinear cutoff to the virtuals is easily achieved by using this result when evaluating the anomalous dimension coefficients. For the reals, the collinear cutoff appears in the direction sampling and the matrix element evaluation.

3.1.7 Direction sampling

Equipped with a collinear cutoff, we can continue by selecting the direction of the next emission. Fortunately, some of the hard work has already been done; the physical distribution over angular phase-space, as given by the emission operators in Eq. (2.1), is found in the first parenthesis of $\mathrm{d}R(E, n, \tau, \bar{\tau})$, Eq. (3.19).

We already selected a dipole to emit from, from the distribution in Eq. (3.23). Now we need to select the direction of the emitted particle according to the physical distribution. However, implementing the collinear cutoff naively would result in holes around the legs

i and j where no radiation is emitted. This creates problems when evaluating the real emission matrix elements using rings and strings, because the other participating dipoles can still radiate close to i and j. For this reason we want the sampled direction to cover the entire 4π solid angle, without any holes. We still need to avoid the radiation diverging at the collinear poles, so we define the distribution:

$$\begin{aligned} \mathrm{d}P_{\text{dir}}(\theta,\phi) &= \frac{\rho\left(\frac{\hat{n}}{(\Lambda\hat{n})^0},\hat{n}_i,\hat{n}_j\right)}{\widetilde{\Omega}_{ij}}\frac{\mathrm{d}\hat{\Omega}}{4\pi\left[(\Lambda\hat{n})^0\right]^2} \\ &= \frac{\rho\left(\hat{n},\hat{n}_i,\hat{n}_j\right)}{\widetilde{\Omega}_{ij}}\frac{\mathrm{d}\hat{\Omega}}{4\pi} = \left(\frac{\omega_{ij}(\hat{n},\hat{n}_i,\hat{n}_j)\Theta_{\text{rap}}+c(1-\Theta_{\text{rap}})}{\widetilde{\Omega}_{ij}}\right)\frac{\mathrm{d}\hat{\Omega}}{4\pi}, \end{aligned} \tag{3.32}$$

where $\rho(\hat{n},\hat{n}_i,\hat{n}_j)$ is the probability density function of the emission direction in the ij-ZMF, and c is a constant defined to take the value of $\omega_{ij}(\hat{n})$ when $\hat{n}$ is at the boundary of the collinear cutoff:

$$c = \omega_{ij}\left(\cos\theta = \frac{1-\lambda/n_i\cdot n_j}{1+\lambda/n_i\cdot n_j}\right) = \frac{(1+\lambda/n_i\cdot n_j)^2}{2\lambda/n_i\cdot n_j}. \tag{3.33}$$

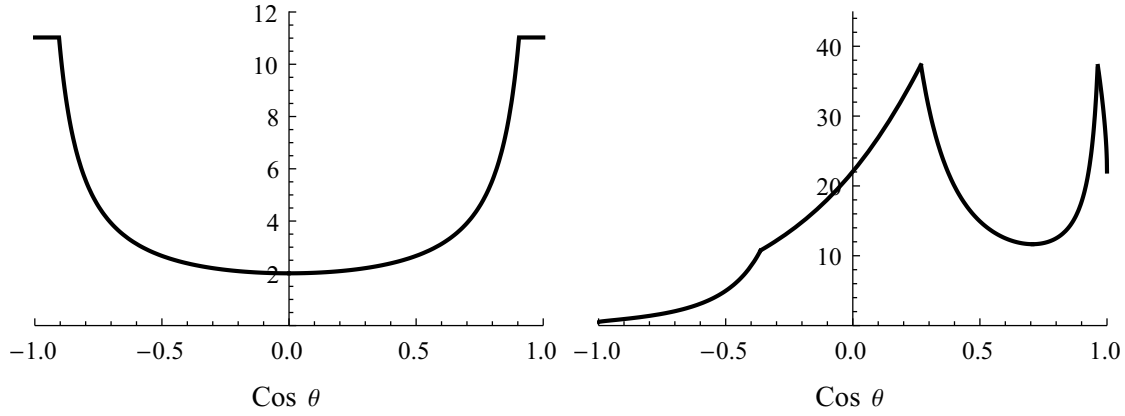
We have also used $\hat{n}_i\cdot\hat{n}_j = n_i\cdot n_j$. Finally, the quantity $\widetilde{\Omega}_{ij}$ is the integral over $\rho(\hat{n},\hat{n}_i,\hat{n}_j)\mathrm{d}\hat{\Omega}/4\pi$. It differs from Ω_{ij}, which we use for the anomalous dimension coefficients, due to the flat regions inside the collinear cutoff cones.

We have defined the distribution in the back-to-back frame of ij as that makes it possible to sample by inversion from it. We can then transform the emission back into the lab frame. However, the constant c picks up a boost factor in the transformation:

$$\begin{aligned} \rho\left(\frac{\hat{n}}{(\Lambda\hat{n})^0},\hat{n}_i,\hat{n}_j\right)\frac{\mathrm{d}\hat{\Omega}}{\left[(\Lambda\hat{n})^0\right]^2} &= \rho(n,n_i,n_j)\mathrm{d}\Omega \\ &= \left(\omega_{ij}(n,n_i,n_j)\Theta_{\text{rap}}+c\left[(\Lambda\hat{n})^0\right]^2(1-\Theta_{\text{rap}})\right)\mathrm{d}\Omega. \end{aligned} \tag{3.34}$$

The radiation patterns in each frame for a slice at fixed ϕ are shown in Fig. 3.3. In the ij-ZMF, radiation increases until it enters the collinear cutoff cones around $(0,0,1)$ and $(0,0,-1)$, where it becomes flat at constant value c. When boosted back into the lab frame, the radiation inside the cutoff cones becomes distorted. The value chosen for the cutoff parameter λ in these plots is much larger than normal, so as to make these features more

prominent.



(a) ij-ZMF: $\rho(\hat{n}, \hat{n}_i, \hat{n}_j)$. **(b)** Lab frame: $\rho(n, n_i, n_j)$.

Figure 3.3: The distributions for the direction sampling of real emissions, which can emit into the collinear cutoff cones, shown for the slice $\phi = 0$. In the lab frame, $n_i = (1, 0, 0, 1)$ and $n_j = (1, 1, 0, 0)$. The collinear cutoff is set to $\lambda = 0.05$.

The radiation distribution in the lab frame is important when we consider the evaluation of the real emission matrix element with rings and strings, in Section 3.1.9. For now, we simply need to sample by inversion from the distribution in the ij-ZMF. We first define two regions, A and B, where $\Theta_{\text{rap}} = 1$ or 0 respectively. We select in which region to emit with probabilities

$$P(A) = \frac{\Omega_A}{\widetilde{\Omega}_{ij}} = \frac{\ln(n_i \cdot n_j/\lambda)}{\widetilde{\Omega}_{ij}} \quad P(B) = \frac{\Omega_B}{\widetilde{\Omega}_{ij}} = \frac{1 + \lambda/n_i \cdot n_j}{\widetilde{\Omega}_{ij}} \tag{3.35}$$

where $\widetilde{\Omega}_{ij} = \Omega_A + \Omega_B$. This selection cancels the $\widetilde{\Omega}_{ij}$ weight introduced when selecting the dipole from Eq. (3.23), and introduces either Ω_A or Ω_B, which will cancel when we sample by inversion from their respective distributions. If the emission is chosen to go into A, the polar angle is sampled from

$$\cos\theta = 1 - \frac{2(\lambda/n_i \cdot n_j)^{\mathcal{R}}}{1 + \lambda/n_i \cdot n_j} \tag{3.36}$$

where $\mathcal{R}$ is a uniform random number in the interval $(0, 1)$. This selection recreates the

central region in Fig. 3.3a. Alternatively, if region B is selected, the distribution

$$\cos\theta = \frac{1 + (2\mathcal{R} - 1)\lambda / n_i \cdot n_j}{1 + \lambda / n_i \cdot n_j} \tag{3.37}$$

generates the flat radiation at the edges of Fig. 3.3a. In either case, ϕ is sampled from a flat distribution in the range $[0, 2\pi]$, and the sampled polar angle is flipped, $\cos\theta \to -\cos\theta$ with probability of a half.

With the direction of the next emission in the ij-ZMF selected, the next step is to transform it back to the lab frame. We use the inverse of the boost defined in Eq. (3.28), which is given by

$$\Lambda = \begin{pmatrix} \sqrt{\frac{2}{n_{ij}}} & \sqrt{\frac{2}{n_{ij}} - 1} & 0 & 0 \\ \sqrt{1 - \frac{n_{ij}}{2}} & 1 & 0 & -\sqrt{1 - \frac{n_{ij}}{2}} \\ 0 & 0 & 1 & 0 \\ \sqrt{\frac{2}{n_{ij}}} - \sqrt{\frac{n_{ij}}{2}} & \sqrt{\frac{2}{n_{ij}} - 1} & 0 & \sqrt{\frac{n_{ij}}{2}} \end{pmatrix} \tag{3.38}$$

where $n_{ij} = n_i \cdot n_j$. In general, n_i and n_j will not be oriented exactly as in Eq. (3.29), and we will need to rotate them back to their lab frame orientation:

$$\begin{aligned} n_i^\mu &= (1, \sin\theta_i \cos\phi_i, \sin\theta_i \sin\phi_i, \cos\theta_i) \\ n_j^\mu &= (1, \sin\theta_j \cos\phi_j, \sin\theta_j \sin\phi_j, \cos\theta_j), \end{aligned} \tag{3.39}$$

which we can do with the rotation matrix given by

$$R = \begin{pmatrix} \frac{(n_{ij}-1)\,\bar{c}_i s_i + \bar{c}_j s_j}{\sqrt{(2-n_{ij})\,n_{ij}}} & \frac{c_j s_i \bar{s}_i - c_i s_j \bar{s}_j}{\sqrt{(2-n_{ij})\,n_{ij}}} & \bar{c}_i s_i \\ \frac{(n_{ij}-1)\,s_i \bar{s}_i + s_j \bar{s}_j}{\sqrt{(2-n_{ij})\,n_{ij}}} & \frac{c_i \bar{c}_j s_j - c_j \bar{c}_i s_i}{\sqrt{(2-n_{ij})\,n_{ij}}} & s_i \bar{s}_i \\ \frac{(n_{ij}-1)\,c_i + c_j}{\sqrt{(2-n_{ij})\,n_{ij}}} & -\frac{\sqrt{-c_i^2 - c_j^2 - 2\,(n_{ij}-1)\,c_i c_j - (n_{ij}-2)\,n_{ij}}}{\sqrt{(2-n_{ij})\,n_{ij}}} & c_i \end{pmatrix} \tag{3.40}$$

where $c_k = \cos\theta_k$, $s_k = \sin\theta_k$, $\bar{c}_k = \cos\phi_k$, and $\bar{s}_k = \sin\phi_k$.

We have now completed the direction sampling of the next emission, which is line 16 in Algorithm 1. This brings us to the second type of veto, indicated with the red line. With the sampled emission four-momentum n in the correct, unrotated lab frame, we can check if the emission was made in the gap region. If it is, the measurement function of the gap veto cross-section, in Eq. (3.2), will set the weight of the whole event to zero, and thus we can terminate it. This veto makes the dedicated mode of `CVolver` very efficient for evaluating this cross-section, as no computational time is spent on events that do not contribute to the observable. When running in event generator mode, the gap veto does not apply (see Section 3.3.2).

Line 17 corresponds to including the weight from the direction sampling. We should summarise all the contributions. First, we selected an emitting dipole according to Eq. (3.23), and the denominator cancelled with the corresponding factor of emission scale proposal in Eq. (3.24). The $\widetilde{\Omega}_{ij}$ in the numerator cancels with the selection of emitting into the A or B region, in Eq. (3.35), and the $\cos\theta$ sampling for either distribution cancels Ω_A or Ω_B respectively. The surviving weight is either $\omega_{ij}(n)$, if in the A region, or $c\left[(\Lambda\hat{n})^0\right]^2$, if in the B region. We will keep these weights in the event until we finish the real emission matrix element evaluation in Section 3.1.9. Another remaining weight is the ξ_{ij} in the numerator of the dipole selection, and we will deal with it when selecting the after-emission colour flows.

3.1.8 Selecting flows after a real emission

With the scale, parent dipole, and direction of the next emission selected, the only missing components are final the colour flows. The probability distribution for selecting the colour flows after an emission is taken from the third term in parentheses in Eq. (3.19):

$$P(\tau,\bar{\tau}) = \frac{\langle\tau|\,\bar{\tau}\rangle}{\langle\sigma|\,\bar{\sigma}\rangle}\frac{\left|\left[\tau\right|\mathbf{T}_i\left|\sigma\right\rangle\left\langle\bar{\sigma}\right|\mathbf{T}_j^\dagger\left|\bar{\tau}\right]\right|}{\xi_{ij}}, \tag{3.41}$$

which has two parts: the ratio of scalar product matrices and the colour charge matrix elements. The colour charge matrix elements, defined in (2.20), are null if the flows $\{\tau,\bar{\tau}\}$ are not accessible from an emission off the flows $\{\sigma,\bar{\sigma}\}$, and otherwise they include factors of λ_i, $\bar{\lambda}_i$, and $1/N_\mathrm{c}$ in the case of singlet emissions. The absolute value is taken to ensure a positive definite probability distribution. The scalar product ratio weighs each pair of flows with respect to their effect in colour space, and can only evaluate to N_c or $1/N_\mathrm{c}$.

Lastly, ξ_{ij}, which normalises the distribution, is defined:

$$\xi_{ij} = \sum_{\{\tau\bar{\tau}\}} \xi_{ij}^{(\tau\sigma\bar{\sigma}\bar{\tau})} = \sum_{\{\tau\bar{\tau}\}} \frac{\langle\tau|\,\bar{\tau}\rangle}{\langle\sigma|\,\bar{\sigma}\rangle} \left| [\tau|\,\mathbf{T}_i\,|\sigma\rangle\,\langle\bar{\sigma}|\,\mathbf{T}_j^\dagger\,|\bar{\tau}] \right|, \tag{3.42}$$

and we used $\xi_{ij}^{(\tau\sigma\bar{\sigma}\bar{\tau})}$ in Section 2.3 to calculate the contribution to colour of an emission from a dipole, ring, or string.

The evaluation of $P(\tau, \bar{\tau})$ is performed as follows. We consider first the emission of two gluons, with no singlets. If i and j are quarks or anti-quarks, there will only be one possible way of emitting gluons, and only one corresponding pair $\{\tau\bar{\tau}\}$. If i and j are a quark or anti-quark and a gluon, there will be two possible ways of attaching the gluon emission, and two corresponding pairs $\{\tau\bar{\tau}\}$, as one of the emissions can attach to either the colour line or anti-colour line of the gluon. If i and j are two gluons, there will be four possible ways of attaching the emissions, and four pairs $\{\tau\bar{\tau}\}$. For each accessible pair of after-emission colour flows $\{\tau\bar{\tau}\}$, there will be two more legs l and k from which the emissions can occur, and which lead to the same colour flows. For example, if i and j are a quark and anti-quark, which are colour connected to anti-quark l and quark k respectively, the flows after the emission will be equivalent whether the amplitude emission attaches to i or l and the conjugate amplitude emission attaches to j or k, see Fig. 2.2 and the text around it. Therefore, for each pair $\{\tau\bar{\tau}\}$ we have four participating legs. Note that these legs are not necessarily distinct; in our example, if i is colour connected to j, then $i = l$ and $j = k$.

Now we consider the possible singlet emissions. If a gluon is emitted from i and a singlet is emitted from j, there can be one or two possible flows τ in the amplitude, depending on whether i is a quark or gluon. There is only one possible $\bar{\tau}$ in the conjugate amplitude, because a singlet emission results in the same flow every time. However, it can attach to any quark or anti-quark, and contribute to the same $\bar{\tau}$. Similarly, if two singlets are emitted, there is only one accessible pair $\{\tau\bar{\tau}\}$ of colour flows, and the singlets could have been emitted off of any quark or anti-quark.

With the whole list of accessible flows, they are weighted and selected according to $P(\tau, \bar{\tau})$. Thus, we end the sampling with a selected pair $\{\tau\bar{\tau}\}$, and a list of possible legs contributing to these colour flows.

Sampling from Eq. (3.41) removes the remaining factor of ξ_{ij} from the weight introduced when selecting the dipole pair with Eq. (3.23). Additionally, we divide by

$(\langle\tau|\,\bar{\tau}\rangle/\langle\sigma|\,\bar{\sigma}\rangle)\left|[\tau|\,\mathbf{T}_i\,|\sigma\rangle\,\langle\bar{\sigma}|\,\mathbf{T}_j^\dagger\,|\bar{\tau}]\right|$ to eliminate the residual contribution from $P(\tau,\bar{\tau})$. We are left with a remaining weight of $\omega_{ij}(n)$ or $c\left[(\Lambda\hat{n})^0\right]^2$, depending on whether the emission was sampled into the A or B region. We conclude the colour flow sampling, lines 18 and 19 in Algorithm 1, and continue with the final step: to evaluate the full real emission matrix element.

3.1.9 Evaluating the emission operator element with rings and strings

We need to reconstruct the full real emission matrix element, as given in Eq. (3.18). For this purpose we define the K functions, which cancel the direction sampling weight we introduced in Section 3.1.7, and evaluate the dipole, ring or string contributing to the colour flows selected in Section 3.1.8. The weight of the event gets multiplied by K to complete the evaluation of the real emission matrix elements, in line 20 of Algorithm 1. We consider first the case of a gluon being emitted in both the amplitude and conjugate amplitude. The corresponding K function is defined

$$K_{ijkl}^{\text{no singlets}} = \frac{1}{2}\frac{\omega_{ij}\Theta_{\text{rap}}^{ij} - \omega_{il}\Theta_{\text{rap}}^{il} - \omega_{kj}\Theta_{\text{rap}}^{kj} + \omega_{kl}\Theta_{\text{rap}}^{kl}}{\widetilde{\omega}_{ij} + \widetilde{\omega}_{il} + \widetilde{\omega}_{kj} + \widetilde{\omega}_{kl}}, \tag{3.43}$$

which requires some explanation. In the numerator we have a sum over four dipole functions, which correspond to the four dipoles that can emit into the same colour flows, as was shown in Fig. 2.5. The legs are those we stored during the selection of colour flows in the previous section. The collinear poles are subtracted according to the rapidity cutoff, and Θ_{rap}^{ij} is simply Θ_{rap}, defined in Eq. (3.27), with respect to legs i and j. This numerator encodes the physical kinematic contribution to the real emission, where we have combined all kinematic dipoles that contribute to this trajectory in colour space, thus aiding in the statistical cancellations across events. Note that, if the emission is not from a ring, then some of the legs will be degenerate, and $w_{ii} = 0$, so the above expression works for emissions from dipoles and strings too. The signs of each dipole are in accordance with the signs within the colour charge matrix elements, $[\tau|\,\mathbf{T}_i\,|\sigma\rangle$ and $\langle\bar{\sigma}|\,\mathbf{T}_j^\dagger\,|\bar{\tau}]$, and the factor of a half arises from $\lambda_i = \bar{\lambda}_j = \sqrt{1/2}$. In the denominator, we have

$$\widetilde{\omega}_{ij} = \omega_{ij}\Theta_{\text{rap}}^{ij} + c\left[(\Lambda\hat{n})^0\right]^2(1 - \Theta_{\text{rap}}^{ij}), \tag{3.44}$$

which is the weight we introduced when performing the direction sampling over the whole solid angle in Section 3.1.7. Since we are now working in the lab frame, the boost factor that appears in front of c needs to be evaluated, which is done through the zeroth component of the relation $\frac{\hat{n}}{(\Lambda\hat{n})^0} = \Lambda^{-1}R^{-1}n$, where Λ^{-1} and R^{-1} are the inverse of the matrices defined in Eqs. (3.38) and (3.40).[1] The purpose of the denominator is to cancel the direction sampling weights across all the dipoles that contribute to this trajectory in colour.

In the case of a dipole emission, where $i = j$ and $k = l$ (see again Fig. 2.5), and the direction was sampled into the A region, where $\Theta_{\text{rap}} = 1$, then $K_{ijkl}^{\text{no singlets}} = 1$. Therefore, we do not introduce any unnecessary weights, and the kinematic factor is included from the importance sampling of the direction. If the emission is from a ring or string, then $K_{ijkl}^{\text{no singlets}}$ may introduce some weight, although the collinear enhancements will cancel explicitly, instead of relying on the statistical cancellation across different Monte Carlo events. Overall, the inclusion of rings and strings proves a significant improvement to the convergence of the Monte Carlo, as shown in [22]. However, we have introduced an inefficiency. If we consider again an emission from a dipole, but this time into the B region, where $\Theta_{\text{rap}} = 0$, then $K_{ijkl}^{\text{no singlets}} = 0$ and the whole event has zero weight. The purpose of the emission cutoff veto, represented by a blue line after line 20 in Algorithm 1, is to veto events where this happens and avoid wasting computational resources.

The root cause of this problem is sampling the direction from a dipole instead of from a ring or string directly. This would also reduce the weights introduced in $K_{ijkl}^{\text{no singlets}}$. At present, implementing importance sampling from a ring or string remains an open avenue for improvement of the `CVolver` algorithm.

We continue by considering the cases with singlet emissions. If a gluon is emitted in the amplitude and a singlet is emitted in the conjugate amplitude, the weighting factor becomes

$$K_{ik}^{\text{one singlet}} = \frac{1}{2N_{\text{c}}} \frac{\sum_{j\in q}\left(-\omega_{ij}\Theta_{\text{rap}}^{ij} + \omega_{kj}\Theta_{\text{rap}}^{kj}\right) + \sum_{j\in\bar{q}}\left(\omega_{ij}\Theta_{\text{rap}}^{ij} - \omega_{kj}\Theta_{\text{rap}}^{kj}\right)}{\sum_{j\in q}\left(\widetilde{\omega}_{ij} + \widetilde{\omega}_{kj}\right) + \sum_{j\in\bar{q}}\left(\widetilde{\omega}_{ij} + \widetilde{\omega}_{kj}\right)}, \tag{3.45}$$

where i is a leg with colour index c_i, which is colour connected to leg k with anti-colour index $\bar{c}_k$. This is equivalent to $K_{ijkl}^{\text{no singlets}}$, but the indices k and l run over all quarks and anti-quarks respectively. The action of the singlet emission operator $\mathbf{s}$ in $\langle\bar{\sigma}|\,\mathbf{T}_j^\dagger\,|\bar{\tau}]$, is

[1] We cannot calculate $\Lambda\hat{n}$ directly because the code for the evaluation of the matrix elements is independent of the direction sampling, and therefore the former does not have access to the coordinates of the emission in the ZMF of the parent legs. Modularisation of the code is essential to keep it organised, readable, and scalable.

accompanied by an explicit factor of $1/N_c$. If the singlet is emitted in the amplitude and the gluon in the conjugate amplitude, the weight $K_{jl}^{\text{one singlet}}$ has the equivalent structure. Finally, if a singlet is emitted on both sides, the weighting factor is given by

$$K^{\text{two singlets}} = \frac{1}{2N_c^2} \frac{2\sum_{i\in q}\left(\sum_{j\in q}\omega_{ij}\Theta_{\text{rap}}^{ij} - \sum_{j\in\bar{q}}\omega_{ij}\Theta_{\text{rap}}^{kj}\right)}{2\sum_{i\in q}\left(\sum_{j\in q}\widetilde{\omega}_{ij} + \sum_{j\in\bar{q}}\widetilde{\omega}_{ij}\right)}, \tag{3.46}$$

where the factors of two in the fraction make explicit that the sums also run for $i \leftrightarrow j$, and there is an explicit suppression of $1/N_c^2$ due to two insertions of the singlet emission operator $\mathbf{s}$.

We have completed our evaluation of the real emission matrix elements, which corresponds to line 20 of Algorithm 1. This step is followed by updating the scalar product matrix, and checking against the colour veto for the third time. We have two sources of colour suppression: first, an explicit factor $1/N_c^q$ from each singlet emission, which we included in our K weights. Therefore, $q = 0, 1$ or 2 for zero, one, or two singlets respectively. The second source is from the update to the scalar product matrix, and we count by how many more swaps the amplitude and conjugate amplitude differ after the emission, which we name r. We count the change in the number of swaps, as we do not want to double count any suppression. The difference in the minimum number of swaps contributes $1/N_c^r$. We already categorised the different types of emissions according to the scalar product matrix in Section 2.3, but we review them here:

- Emission from a dipole: Γ contribution. No more swaps induced, $q+r = 0+0 = 0$.
- Emission from a string: Σ contribution. No more swaps induced, $q+r = 0+0 = 0$.
- Emission from a ring: Σ contribution. Two more swaps induced, $q+r = 0+2 = 2$. Alternatively, no more swaps induced, $q+r = 0+0 = 0$.
- One singlet and one gluon emission: ρ contribution. One swap induced, $q+r = 1+1 = 2$.
- Two singlets emitted: ρ contribution. No more swaps induced, $q+r = 2+0 = 2$.

where we notice that $q+r = 0$ or 2, and crucially, the sum cannot be negative. Therefore, a real emission cannot make the evolution less colour suppressed. Moreover, although some ring emissions and all string emissions maintain the same level of suppression in the evolution, they are still a Σ contribution, and are therefore subleading colour effects.

At this stage we have accrued a colour suppression $1/N_c^T$, where $T = p + (s + t) + (q + r)$. As before, if $T > m$ we terminate the event, where m was an input parameter that we choose as the maximum colour suppression we intend to calculate to. This veto is fundamental in keeping the evolution sufficiently close to the colour diagonal, since with a growing number of emissions there are increased chances of sampling subleading colour. Furthermore, once a level of suppression is attained it cannot be undone by any mechanism in the evolution. Without the colour veto, the results shown in Chapters 4 and 5 would not have been possible to generate.

The final part of Algorithm 1, lines 22, 23 and 24 correspond to setting up the variables necessary to evaluate the next virtual evolution: set the new scale to evolve from, set the new colour flows to evolve from as the colour flows after the emission, and increasing the number of emissions by one. The loop of emissions then continues until the evolution reaches the evolution cutoff ρ, the maximum number of emissions is reached, or one of the vetos triggers.

As the number of emissions increases, the complexity of the steps described in this chapter can grow very large. In particular, the number of colour flows grows factorially with the number of emissions, which can quickly become computationally intractable. The algorithm in `CVolver` is only feasible because it tracks systematically every possible step taken in colour space, and prioritises the trajectories that contribute most to the cross-section. To illustrate this point, and to provide a practical example of everything we have discussed thus far, in the next section we will follow the whole algorithm, step by step, for a single `CVolver` event with four emissions.

3.2 An example: one event, step by step

The previous section covered every step of Algorithm 1 in detail, and the implementation into the `CVolver` Monte Carlo. Now we consider an actual event, print out every weight and selection made, and analyse its evolution.[2] We evolve a $q\bar{q} \to q\bar{q}$ hard process, and we have set the maximum colour suppression to $m = 5$. The output is shown in Table 3.1, and the structure follows the order of the previous section and of Algorithm 1: The first three rows correspond to setting the hard process and the evolution cutoff. It is then followed by the emissions loop, which contains six steps: the selection of virtual evolution colour flows, selection of the next emission scale, evaluation of the virtual matrix element, sampling the emission direction, selection of the flows after real emission, and evaluation of the real emission matrix element. For the first run through the loop, which corresponds to the 1^{st} emission, the subsections that describe each step are given. For the subsequent emissions, the emission number is given instead. We will overview the evolution, and point out interesting features.

The first two steps are selecting the hard process kinematics and the colour flows. The kinematics are set according to Eq. (3.7) with $\theta = \pi/6$, and the colour flows are set to $|10\rangle\langle 01|$. This is one of the kinematic configurations that we consider in Chapters 4 and 5, and in fact this event contributes to Fig. 4.11. We are evolving an interference contribution, which is very badly described without including all subleading colour effects, as discussed next chapter. This example will provide some insight into this result. For now, we note that the evolution starts at colour suppression $1/N_c^T$, where $T = p = 1$, and we have picked up this suppression from the scalar product of the interference, $\langle 10|01\rangle = N_c$, relative to the highest order in colour contributing to this process, which is $\langle 01|01\rangle = \langle 10|10\rangle = N_c^2$.

The veto scale, or evolution cutoff of the event is chosen to be $\rho = 0.15$, and a corresponding weight applied. The virtual evolution proceeds uneventfully; no swaps are selected on either the amplitude nor conjugate amplitude, and the scale of the first emission is chosen to be $E_1 = 0.829$, which is above the veto scale and below the maximum number of emissions, so the emission is allowed. The direction is sampled into region A from parents a and b, and we note that no weight needs to be applied as the sampling distributions follow the physical real emission distribution. Any weight from other dipole

[2] This event can be recreated with the version $1139 : 93f55d4193ca$ of `CVolver`, and running the command `dijetVetoHistograms -d 2 -n 20000000 -r 8 -s 1 -rc 0.1 -nbins 40 -lcprime 1 -hardprocess 4 -initialFlows 10 01 -emit2max 0 -l 0.01 -inclusive 0 -no-singlets 0 -beta 4 -cutoff rapidity -leadingNflows 0 -beta-switch 1 -dSup 1 -cone 4 -dipoleangle 6 -N 3 -colour-limit 5 -virtual-weights-in-proposal 0 -nMin 0 -colourOrderMin 0 -coulomb 0`. It is the first event for this seed to fail the colour veto during the $4^{th}$ emission, and it can be found with a corresponding `assert()` line in the code.

contributions will be included in the K factors defined in Section 3.1.9. Indeed, the flows after real emission are selected to be from two singlet emissions, which will involve multiple kinematic dipole contributions. First we check that the weight applied corresponds with dividing out the numerator of the real emission colour flows distribution in Eq. 3.1.8:

$$\begin{aligned}\frac{\langle\tau|\,\bar{\tau}\rangle}{\langle\sigma|\,\bar{\sigma}\rangle}\left|[\tau|\,\mathbf{T}_a\,|\sigma\rangle\,\langle\bar{\sigma}|\,\mathbf{T}_b^\dagger\,|\bar{\tau}]\right| &= \frac{\langle 102|\,012\rangle}{\langle 10|\,01\rangle}\left|[102|\,\mathbf{T}_a\,|10\rangle\,\langle 01|\,\mathbf{T}_b^\dagger\,|012]\right| \\ &= N_\mathrm{c}\bar{\lambda}_a\lambda_b\frac{1}{N_\mathrm{c}^2} = \frac{1}{6}.\end{aligned} \tag{3.47}$$

which agrees with the weight applied $w \to w \times 6$. Since singlets can be emitted from any quark or anti-quark, there are six distinct dipoles participating in $K^{\text{two singlets}}$: ab, ac, ad, bc, bd, cd. It is instructive to consider some of them: the coordinates of the emission were sampled to be close in angle to n_b, which explains how $\omega_{ab} \simeq 70, \omega_{bc} \simeq 68$ are enhanced relative to $\omega_{ac} \simeq 0.03$. Furthermore, ω_{ab} and ω_{bc} appear with opposite signs in the numerator of $K^{\text{two singlets}}$, due to ab being a colour and anti-colour line, and bc being two colour lines. These enhancements cancel and the overall factor is small: $K^{\text{two singlets}} = 0.020$. If we were not including all contributing dipoles then only $\omega_{ab} \simeq 70$ would survive in the event, since ab was the selected dipole originally, and it would require statistical cancellation with the other dipoles across different events. It is evident that the inclusion of all contributing dipoles in the K factors is essential to the convergence of the Monte Carlo.

We should also pay attention to the accumulated colour suppression at this point. The two singlet emissions increase T by $q = 2$, but the scalar product remains at the same suppression. This corresponds to the last case listed at the end of Section 3.1.9. Therefore, $T < m = 5$ and the evolution continues.

The next interesting feature occurs at the selection of colour flows after the second emission. The selected flows, $|1320\rangle\langle 3120|$, correspond to an emission in the amplitude from colour line 1 or anti-colour line $\bar{0}$, and in the conjugate amplitude from colour line 0 or anti-colour line $\bar{0}$, which is a string emission as shown in Fig 2.5. Moreover, every possible non-singlet emission from these set of flows is a string contribution which are subleading colour Σ terms. Therefore, if no subleading colour emissions were allowed, no emissions could occur whatsoever when evolving an interference contribution. This explains why strict leading colour evolution completely fails at describing the evolution of interferences in Chapter 4. Again, the values of the contributing dipoles are noteworthy: $\omega_{ab} \simeq 81, \omega_{ca} \simeq 91$, and $\omega_{cb} \simeq 8$. The first two contribute with opposing signs, and

the resulting weight is kept under control, $K^{\text{no singlets}_{abcb}} \simeq 0.457$. Also, string emissions, although subleading colour effects, do not contribute to any further colour suppression, and $T = 3$ remains.

The following virtual evolution, starting the loop for the third emission, has dramatic effects; two swaps are selected in the amplitude, and one in the conjugate amplitude. The low probability of this occurring is associated with a large weight $w \to w \times 338$. But the remarkable effect is the swaps reduce the scalar product matrix suppression: one of the amplitude swaps brings the flows to the colour diagonal (where the event has never been so far as it started from an interference) and the remaining swaps on each side match each other. Therefore, the evolution is now at the colour diagonal, $|0123\rangle\langle 0123|$. However, it has accumulated some colour suppression on its way: from this step, the event gains $s + t = 3 - 1$, corresponding to an explicit factor of $1/N_\text{c}^3$ from the Sudakov expansion, and -1 due to gaining a factor of N_c in the scalar product. This brings the total colour suppression to $T = 5$, which as the limit of what is allowed. Also, the Sudakov matrix element decreases the overall weight significantly: $w \to w \times (-1.33 \times 10^{-8})$, which justifies and counteracts the number of swaps distribution allocating a small probability to this contribution.

The event then proceeds, and survives the third emission: it is a dipole emission, and therefore a leading colour contribution. The following virtual evolution selects no further swaps. However, the event then emits another singlet in the fourth emission loop, which brings T above m, and the event is terminated. At the end, it had accumulated a factor of $1/N_\text{c}^7$, which made it too suppressed to be worth the computational resources, and it was terminated and discarded.

We have showcased many fundamental and interesting features of the amplitude level evolution in `CVolver`. Not only can interferences be evolved, but their full colour behaviour cannot be captured whatsoever without subleading colour effects. We have also shown how Algorithm 1 minimises large weights by importance sampling when possible, and how ring and string structures are not only essential to breaking down the evolution in colour, but they provide systematic cancellation of collinear enhancements. Most importantly, the algorithm described here is able to track every order in $1/N_\text{c}$ at each step in the soft evolution. These are the ingredients that make the results in Chapters 4 and 5 possible.

Designing a Monte Carlo that explores efficiently all possible colour trajectories is a challenging task. The colour veto is necessary to avoid the evolution drifting too subleading in colour, especially for higher multiplicities. Furthermore, throughout the algorithm

we described many possible settings and parameters to steer the evolution. Next section we summarise most of the options currently implemented in `CVolver`, including those we use to improve the convergence of our results.

Step within evolution	Sampled quantities and weights	Colour flows	Colour suppression $T=p+(s+t)+(q+r)$
Hard process: kinematics Section 3.1.1	$n_a(x,y,z) = (0.529, -0.700, -0.479)$ $n_c(x,y,z) = (0.879, -0.353, -0.321)$ $n_b = -n_a$, $n_d = -n_c$ $w = 1$	N/A	$T = 0 = 0 + (0+0)+(0+0)$
Hard process: colour flows Section 3.1.1	Flows set manually $w \to w \times 1$	$\|10\rangle\langle 01\|$	$T = 1 = 1 + (0+0)+(0+0)$
Evolution cutoff selection Section 3.1.2	$\rho = 0.150$ $w \to w \times 0.345$	$\|10\rangle\langle 01\|$	$T = 1 = 1 + (0+0)+(0+0)$
Evolution matrix: next flows Section 3.1.3	$d' = \overline{d}' = 0$ $w \to w \times 1.778$	$\|10\rangle\langle 01\|$	$T = 1 = 1 + (0+0)+(0+0)$
Next emission scale Sections 3.1.4 and 3.1.5	$E_1 = 0.829$ Dipole selected: a, b Case 1 $w \to w \times 2.485$	$\|10\rangle\langle 01\|$	$T = 1 = 1 + (0+0)+(0+0)$
Evolution matrix: evaluation Section 3.1.5	$w \to w \times 0.857$	$\|10\rangle\langle 01\|$	$T = 1 = 1 + (0+0)+(0+0)$
Emission matrix: direction sampling Section 3.1.7	Region A $n_1(x,y,z) = (-0.604, 0.726, 0.329)$ $w \to w \times 1$	$\|10\rangle\langle 01\|$	$T = 1 = 1 + (0+0)+(0+0)$
Emission matrix: next flows Section 3.1.8	Two singlets $w \to w \times 6$	$\|102\rangle\langle 012\|$	$T = 3 = 1 + (0+0)+(2+0)$

Continued from previous page

Step within evolution	Sampled quantities and weights	Colour flows	Colour suppression $T=p+(s+t)+(q+r)$
Emission matrix: evaluation Section 3.1.9	Six participating dipoles $w \to w \times 0.020$	$\|102\rangle\langle 012\|$	$T = 3 = 1 + (0+0)+(2+0)$
Evolution matrix: next flows (2nd emission)	$d' = \bar{d}' = 0$ $w \to w \times 1.778$	$\|102\rangle\langle 012\|$	$T = 3 = 1 + (0+0)+(2+0)$
Next emission scale (2nd emission)	$E_2 = 0.735$ Dipole selected: a, b Case 1 $w \to w \times 4.635$	$\|102\rangle\langle 012\|$	$T = 3 = 1 + (0+0)+(2+0)$
Evolution matrix: evaluation (2nd emission)	$w \to w \times 0.905$	$\|102\rangle\langle 012\|$	$T = 3 = 1 + (0+0)+(2+0)$
Emission matrix: direction sampling (2nd emission)	Region A $n_2(x, y, z) = (0.547, -0.592, -0.592)$ $w \to w \times 1$	$\|102\rangle\langle 012\|$	$T = 3 = 1 + (0+0)+(2+0)$
Emission matrix: next flows (2nd emission)	String emission $w \to w \times 0.666$	$\|1320\rangle\langle 3120\|$	$T = 3 = 1 + (0+0)+(2+0)$
Emission matrix: evaluation (2nd emission)	Three participating dipoles $w \to w \times 0.457$	$\|1320\rangle\langle 3120\|$	$T = 3 = 1 + (0+0)+(2+0)$
Evolution matrix: next flows (3rd emission)	$d' = 2, \bar{d}' = 1$ $w \to w \times 338$	$\|0123\rangle\langle 0123\|$	$T = 5 = 1 + (3-1)+(2+0)$
Next emission scale (3rd emission)	$E_3 = 0.299$ Dipole selected: a, b Case 1 $w \to w \times 7.66$	$\|0123\rangle\langle 0123\|$	$T = 5 = 1 + (3-1)+(2+0)$

Continued from previous page

Step within evolution	Sampled quantities and weights	Colour flows	Colour suppression $T=p+(s+t)+(q+r)$
Evolution matrix: evaluation (3rd emission)	$w \to w \times (-1.33 \times 10^{-8})$	$\|0123\rangle\langle 0123\|$	$T = 5 = 1 + (3-1)+(2+0)$
Emission matrix: direction sampling (3rd emission)	Region A $n_3(x, y, z) = (-0.487, 0.859, 0.154)$ $w \to w \times 1$	$\|0123\rangle\langle 0123\|$	$T = 5 = 1 + (3-1)+(2+0)$
Emission matrix: next flows (3rd emission)	Dipole emission $w \to w \times 0.666$	$\|41230\rangle\langle 41230\|$	$T = 5 = 1 + (3-1)+(2+0)$
Emission matrix: evaluation (3rd emission)	One participating dipole $w \to w \times 0.5$	$\|41230\rangle\langle 41230\|$	$T = 5 = 1 + (3-1)+(2+0)$
Evolution matrix: next flows (4th emission)	$d' = \overline{d}' = 0$ $w \to w \times 2.09$	$\|41230\rangle\langle 41230\|$	$T = 5 = 1 + (3-1)+(2+0)$
Next emission scale (4th emission)	$E_4 = 0.229$ Dipole selected: a, 1 Case 1 $w \to w \times 5.27$	$\|41230\rangle\langle 41230\|$	$T = 5 = 1 + (3-1)+(2+0)$
Evolution matrix: evaluation (4th emission)	$w \to w \times 0.703$	$\|41230\rangle\langle 41230\|$	$T = 5 = 1 + (3-1)+(2+0)$
Emission matrix: direction sampling (4th emission)	Region B $n_4(x, y, z) = (-0.890, 0.238, 0.388)$ $w \to w \times 1$	$\|41230\rangle\langle 41230\|$	$T = 5 = 1 + (3-1)+(2+0)$
Emission matrix: next flows (4th emission)	One singlet $w \to w \times 6$	$\|412305\rangle\langle 412350\|$	$T = 7 = 1 + (3-1)+(3+1)$

Continued from previous page

Step within evolution	Sampled quantities and weights	Colour flows	Colour suppression $T=p+(s+t)+(q+r)$
Emission matrix: evaluation (4^{th} emission)	Five participating dipoles $w \to w \times -0.005$	$\lvert 412305\rangle\langle 412350\rvert$	$T = 7 = 1 + (3-1)+(3+1)$
Event termination	Colour veto failed $T > m$ $w \to w \times 0$	$\lvert 412305\rangle\langle 412350\rvert$	$T = 7 = 1 + (3-1)+(3+1)$

Table 3.2: An example of all the steps taken by `CVolver` for one event. The hard process is $q\bar{q} \to q\bar{q}$, the collinear cutoff $\lambda = 0.01$, the Sudakov operator expansion is taken at order $d = 2$, the maximum number of emissions is $n_{\text{max}} = 8$, and the colour suppression limit is $m = 5$. The variable w is the weight of the event. The hard process partons are called a, b for the incoming, and c, d for the outgoing. The subsequent emissions are numbered.

3.3 Settings and algorithm modifications

Throughout this chapter we have described some of the input parameters to the evolution in `CVolver` in the order they appear in Algorithm 1. We list here most of the options currently implemented. The parameters listed here are enough to recreate all the results presented in Chapter 4 and 5. These parameters relate to the basic settings of the evolution:

- (`-r`): Maximum number of emissions allowed n_{max}.
- (`-l`): Collinear cutoff parameter λ, used in Eq. (3.27).
- (`-rc`): Lowest veto scale to evolve down to ρ_c, used in Eq. (3.9).
- (`-N`): Number of colours N_{c}.
- (`-s`): Sets the seed for the random number generator in the Monte Carlo.
- (`-n`): Sets the number of events requested. The Monte Carlo will run until this number of events is reached. It does not count events which get vetoed, only events with non-zero weight.

The last three were not mentioned in Section 3.1. The parameter `-N` sets the value of N_{c} for the whole calculation, making it possible to suppress or enhance the contribution of subleading colour. The full colour gap veto cross-section for different values of N_{c} is presented in Fig. 4.2. The following parameters control how subleading colour is handled in the evolution:

- (`-d`): Maximum number of virtual swaps allowed per insertion of the Sudakov operator, used in Eq. (3.13).
- (`-lcprime`): Enable the LC′approximation, which exponentiates the ρ contributions in the expansion of the Sudakov operator, given by Eq. (2.35).
- (`-leadingNflows`): Enforce leading colour evolution in the reals. Therefore, no ring, string, or singlet emissions are allowed.
- (`-colour-limit`): Sets m, where $1/N_{\text{c}}^m$ is the maximum amount of colour suppression allowed before an event is vetoed.
- (`-coulomb`): Includes $i\pi$ factors in the anomalous dimension.

To achieve strict leading colour evolution, the `-leadingNflows` option can be used in combination with setting the maximum number of virtual swaps d to zero, and disabling the LC′ approximation with `-lcprime false`. The different approximations to subleading colour shown in Chapter 4 are achieved with different combinations of the above settings, with the exception of L3 (LC′, dipole and singlet emissions allowed, but no ring or strings emissions). For this type of evolution it is necessary to make a simple modification to the code that implements the real emission colour flows distribution in Eq. (3.41), so only singlets or dipole emissions can be selected.

These parameters control the hard process and the gap region:

- (`-cone`): Defines the gap region. In dedicated mode, if an event is emitted into the gap the event is vetoed. Discussed in Section 3.1.1.
- (`-hardprocess`): Choose hard scattering process, options are shown in Table 3.1.
- (`-initialFlows`): Choose hard process colour flows, discussed in Section 3.1.1.
- (`-dipoleangle`) and (`-jet-rapidities`): Choose hard process kinematics, discussed in Section 3.1.1.

It is very simple to implement in the code more hard scattering processes, kinematics or gap definitions.

3.3.1 The sampling distributions

We continue with the parameters that shift the sampling distributions:

- (`-alpha`): Sets α, which enhances the sampling of events with small veto scales. Used in Eq. (3.9).
- (`-dSup`): Sets s, which suppresses the sampling of virtual swaps. Used in Eq. (3.13).
- (`-beta`): Sets β, which shifts the sampling of emission scales to lower values, and therefore reduces the average number of emissions per event. Used in Eq. (3.24).
- (`-beta-switch`): Sets at which multiplicity $\Theta_\beta(n)$ switches to 1, activating the β parameter and reducing the rate of emissions. Used in Eq. (3.24).

- (`-virtual-weights-in-proposal`): Sets $\Theta_{W_V} = 0$ or 1, which pushes selection weight from the colour flows after virtual evolution into the emission scale sampling. Used in Eq. (3.24).

These parameters do not affect the physical predictions of the evolution, as shifts in the sampling distributions are exactly compensated by corresponding weights. Nevertheless, modifying them impacts statistical convergence. In the most pathological cases, the sampling could be avoiding an entire region of phase-space, neglecting its contribution. For example, if the β parameter is too small, the average event will have very high multiplicity. This can result in close to zero events with 0 emissions below some energy scale, completely missing this contribution to the cross-section.

In Table 3.3 we provide the choice for sampling parameters that have been used for the results presented in the following chapters. In the cases of $Z \to q\bar{q}$ and $H \to gg$, these choices were tested in depth and they were found to produce the smoothest results. Increasing the β parameter tends to improve the convergence of low-multiplicity results at the cost of the high multiplicities. This is expected, as not only do they reduce the proportion of events sampled with high multiplicity, they also accrue a weight proportional to β^n, where n is the number of emissions. The beta switch parameter $\Theta_\beta(n)$ was implemented to solve this issue: by activating the β parameter at a specific multiplicity, it lets events naturally get to the desired multiplicity without accruing unnecessary weight. Therefore, these two parameters can be used in combination to enhance the statistics for different multiplicity events. The total cross-section can then be reconstructed by combining runs with different sampling parameters. This proved a very useful method for generating smooth results.

Hard process	s	β	$\Theta_\beta(n)$	Θ_{W_V}
$Z \to q\bar{q}$ (0 - 4) emissions	5	4	0	1
$Z \to q\bar{q}$ (5 - 10) emissions	5	4	4	1
$H \to gg$	5	4	0	1
$q\bar{q} \to q\bar{q}$	1	3	0	0
$qg \to qg$	1	3	0	0
$gg \to gg$	1	4	0	0

Table 3.3: Sampling parameters used for the full colour results shown in Chapters 4 and 5. In all cases, α was chosen to be 1.

For the $2 \to 2$ processes, the sampling parameters were studied in less depth, so it is

unlikely these are the optimal choices. However, they were good enough to produce the results in the following chapters. By tweaking these, and in particular $\Theta_\beta(n)$, it would be possible to generate smoother results for higher multiplicities.

We note as well that, on top of accessing different phase-space regions and reducing large weights, another factor to consider is the impact on the run time of the Monte Carlo. For example, decreasing β results in more events with higher multiplicities, which take more time to compute. For this reason α is set to 1 for all our results, as more events with low evolution cutoffs significantly increases the average time taken per event, and we did not explore low enough values of ρ to require shifting the evolution cutoff distribution.

Finally, the best choices for sampling parameters also depend on the amount of subleading colour. All the discussion above pertains to full colour evolution, which is the computationally challenging task. For leading colour evolution $\beta = 1$ is the best choice, and the other parameters will have no effect. Similarly, smooth leading colour results can be very easily generated from `CVolver` with $\sim 10^7$ events, while full colour typically requires $\sim 10^{11}$ events.

3.3.2 Event generator mode

So far in this chapter we have considered the "dedicated" mode of `CVolver`, which is optimised for the gaps-between-jets observable. However, `CVolver` is more than a resummation tool for a specific calculation, and it can be run as a partonic event generator. The "event-generator" mode of `CVolver` is used extensively for the results presented in Chapter 5, and it requires a few changes to Algorithm 1.

The goal is to dress some hard-process with soft QCD radiation and obtain a weight w, and a set of kinematics for the emitted gluons $\{k\}_n$. We also store the colour suppression $1/N_c^T$, in order to analyse the subleading colour effects of any observable. Then, a measurement function $u(\{k\}_n)$ can be applied. The changes to the algorithm required are relatively simple: first, the IR cutoff of the evolution μ is selected as an input parameter, and all events must evolve down to μ since the inclusivity of the observable below some scale is a priori unknown. This means the steps in Section 3.1.2 are skipped. The evolution then proceeds as normal, except that the event is not terminated if some kinematic condition is met. Overall, in Algorithm 1 the lines 4, 5, and the gap veto between lines 16 and 17 are skipped.

We note that in the current implementation recoil effects are not accounted for, and

therefore conservation of four-momentum is neglected. Presently, we aim to evaluate observables at LL accuracy, where strong ordering of the evolution variable can be assumed and each emission can be treated independently. To improve the accuracy to NLL and beyond, a recoil prescription will be needed.

It is possible to recreate the gaps-between-jets observable, as a cross-check between the two modes of `CVolver`. This observable vetoes events with emissions within some fixed angular region above an inclusivity scale ρ, so the lowest possible value of this inclusivity scale must be found, ρ_{min}, such that the event would still contribute. The particles within the gap region are selected, and then ρ_{min} is given by

$$\rho_{\text{min}} = \max\{E_i, i \in \Omega_{\text{gap}}\}. \tag{3.48}$$

Therefore, each event contributes to $\Sigma(\rho)$ for any $\rho > \rho_{\text{min}}$; the distribution is filled cumulatively. Results showing agreement between the two modes are shown in Chapters 4 and 5.

The parameters that control the event generator mode of `CVolver` are listed here:

- (`-ircutoff`): Select infrared cutoff μ to evolve down to in event generator mode.
- (`-inclusive`): Disable final virtual evolution when number of emissions reaches n_{max}.
- (`-nMin`): Sets the minimum number of emissions an event must have to store its data.
- (`-colourOrderMin`): Sets the minimum colour suppression an event must have to store its data.

The `-inclusive` setting is required to account for further emissions that could occur past n_{max}. In event-generator mode an event with infinite emissions will still contribute to $\sigma_{n=0}(\rho)$ of the gaps-between-jets observable if the first particle was emitted in-gap. For that purpose, removing the final virtual evolution for the final emission makes that cross-section inclusive over any number of extra emissions, and provides a correction over simply evolving down to μ when n_{max} is reached. Algorithmically, inclusive mode always follows Cases 1 and 2 in Section 3.1.5, even for $n = n_{\text{max}}$. For the final emission it also skips the virtual evolution, lines 7, 8, 11, 12 in Algorithm 1, and then terminates the event. However, for the results presented in Chapter 5 the maximum number of emissions is always chosen

to be larger than what is numerically needed, so inclusive mode does not make a significant impact.

The last two options listed, `-nMin` and `-colourOrderMin`, relate to the data storage component of `CVolver`. Since the soft evolution in event-generator mode is observable independent, it is computationally sensible to store the generated data, which can then be reanalysed without significant computing power needed. `CVolver` stores its events using the `HDF5` data format [107, 108], and for each event it stores the hard process information, the weight, the colour suppression, and the kinematics of the emissions. The data storage module of `CVolver` was implemented by the author of this thesis, and it links to the `myStatistics` package to produce histograms [109]. Due to the large number of events needed for statistical convergence, the amount of disk space required can be very large. It strongly depends on the average number of emissions per event, but the data for the evolution of a single hard process kinematic configuration, from a specific set of initial flows, occupies ~ 700GB. The total amount of data generated for the studies presented in Chapter 5 occupies ~ 25TB. The options `-nMin` and `-colourOrderMin` are used to store only the most statistically significant events. For example, 0 emissions at leading colour can be easily produced with a few thousand events, and storing any more would be unnecessary.

Chapter 4

Exact colour evolution for jet observables

In the previous chapter we described how the amplitude level evolution defined in Chapter 2, based on work done in [20], can be implemented as a Monte Carlo parton shower, `CVolver`. We discussed how `CVolver` is able to track all factors of $1/N_c$ that appear in the evolution, and how the sampling is steered in order to maximise computing efficiency. Some initial results from `CVolver` have been published in previous publications [23, 24], but the results presented here correspond to the most up-to-date and efficient implementation. We use the dedicated mode, defined in Section 3.1, to perform a systematic and comprehensive analysis of subleading colour corrections in the gaps-between-jets observable. This chapter is based on a publication [103] which was produced in collaboration with J. R. Forshaw and S. Plätzer.

The remainder of this chapter is organised as follows. Section 4.1 presents results on primary dijet production from a colour singlet initial state, as might occur in $Z \to q\bar{q}$ or $H \to gg$. We focus our attention on the jet veto cross-section for fixed kinematics of the jets, i.e. we veto the production of additional gluons in some fixed angular region if their energy exceeds some threshold. Our goal is to provide a systematic study of sub-leading colour effects in a clear and controlled manner. In Section 4.2 we move on to study dijet production in hadron-hadron collisions, again for fixed parton kinematics. Finally, we look at the colour singlet production of $q\bar{q}q\bar{q}$, as might occur following the hadronic decay of two Z bosons.

4.1 Colour singlet production of dijets

We start with the production of a pair of jets from a colour singlet initial state, such as might be produced in $Z \to q\bar{q}$ or $H \to gg$. We fix the hard process kinematics such that the two jets are produced back-to-back, each with energy Q. Additional radiation is vetoed if it falls in the interjet region, which is defined by cones centred on each jet with an opening angle of $\pi/2$, and has energy $E > Q_0$, as illustrated in Fig. 4.1. The relevant large logarithm is $\ln(1/\rho)$ where $\rho = Q_0/Q$. For this process the hard process density matrix is simply $H = N_c$, corresponding to a single colour flow.

In Fig. 4.2 we show how `CVolver` is able to compute the terms in the $1/N_c$ expansion. The figure shows the cross-section (normalized to unity at $\rho = 1$) for different values of N_c (these are the different line types). The different colours show the contributions from each colour order. We could present results strictly in the expansion of $1/N_c$ but since it is very easy to sum the colour diagonal virtual corrections to all orders (our LC$'$ approximation) we choose to do so. Specifically, the black curves are computed at full colour, the red

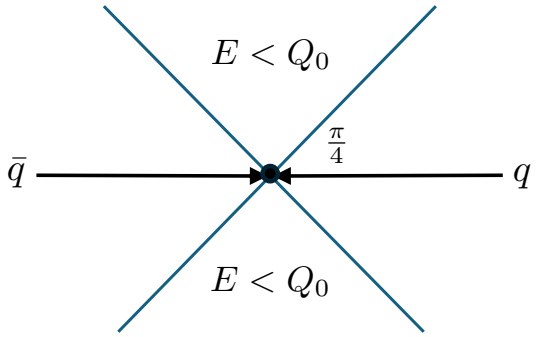


Figure 4.1: Two partons (here a quark and anti-quark) are produced back-to-back and radiation is vetoed in between the corresponding jets.

curves are the LC′ contribution, the blue curves are the $1/N_c^2$ suppressed contributions, the orange are the $1/N_c^4$ suppressed contributions, etc. The solid green curve indicates that the $1/N_c^6$ contribution is suppressed by 3 orders of magnitude, which is as one would naively anticipate. The figure also illustrates the effect of changing the numerical value of N_c and, as expected, we see an increase in the size of the N_c suppressed terms as N_c is reduced. In what follows we shall have much more to say about the nature of the sub-leading colour contributions. For now we simply note that the subleading colour contributions to the LC′ approximation are substantial ($\approx 20\%$ for $N_c = 3$ at $\rho = 10^{-2}$).

Fig. 4.3 shows the veto cross-section broken down by multiplicity and colour accuracy. The line-types used in this plot will be used consistently throughout the remainder of the chapter for processes involving any number of quarks or anti-quarks. They are defined as follows.

- Short-dashed: the strictly leading colour result. We will refer to this as our L1 result.
- Dotted: the LC′ result with the leading colour approximation in the real emissions. We will refer to this as our L2 result.
- Long-dashed: the LC′ result but with the real emissions computed by also including singlet emissions. This approximation should be closest to that of a standard parton shower algorithm. We will refer to this as our L3 result. We emphasise that the analogy to standard parton showers refers only to their treatment of colour; many effects, such as the running of the coupling and recoil, among others, are not included and can significantly alter the picture.
- Dash-dotted: the LC′ result with a full colour treatment of the real emissions. This is the first approximation to include dipoles, rings and strings. It is our L4 result.

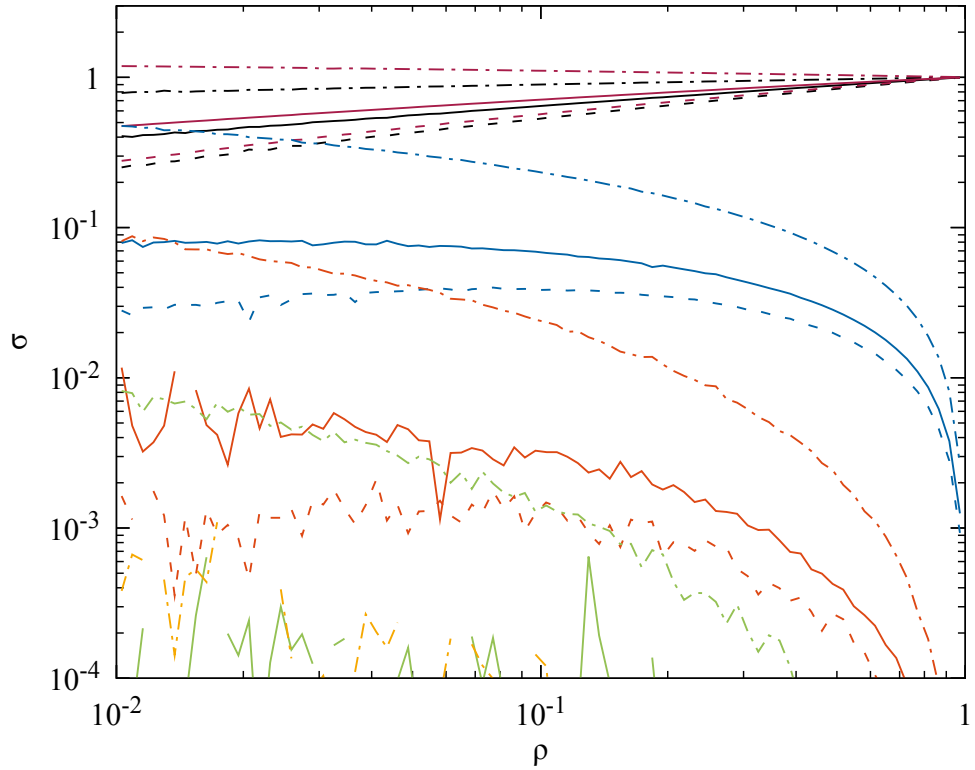


Figure 4.2: The veto cross-section as a function of veto scale for $Z \to q\bar{q}$. Solid: $N_c = 3$, Dash-dotted: $N_c = \sqrt{2}$, Dashed: $N_c = 4$. Black: Full colour, Red: LC′, Blue: NNLC′, Orange: N^{4}LC′, Green: N^{6}LC′, Yellow: N^{8}LC′. The blue and green contributions are negative.

- Solid: the full colour result. Here we go beyond the LC′ approximation for the virtual corrections, i.e. we go beyond $d = 0$ in Eq. (2.35). This involves substantially greater computational effort than the other approximations and is our L5 result.

A striking observation here is that the L4 and L5 results agree at the percent level. This is in accord with the observations made by Hatta and Ueda in [110] and PanScales in [111], if we assume that our L4 result is approximately equal to the mean field approximation of [110] and the two schemes presented in [111]. A corollary to this is that we can safely use the $d = 0$ LC′ approximation in the expansion of the virtual corrections. In other words, $N_c/2 \to C_F$ in the diagonal part of the virtual corrections combined with a full colour treatment of the real emissions is sufficient to capture sub-leading colour effects in $Z \to q\bar{q}$. As we shall see, this feature does not extend to other processes. Note also that the zero emission curves for L2–L4 are exactly degenerate since the subsequent enhancements only impact upon the real emissions.

Fig. 4.4 shows three control plots to illustrate that `CVolver` is working well. The first

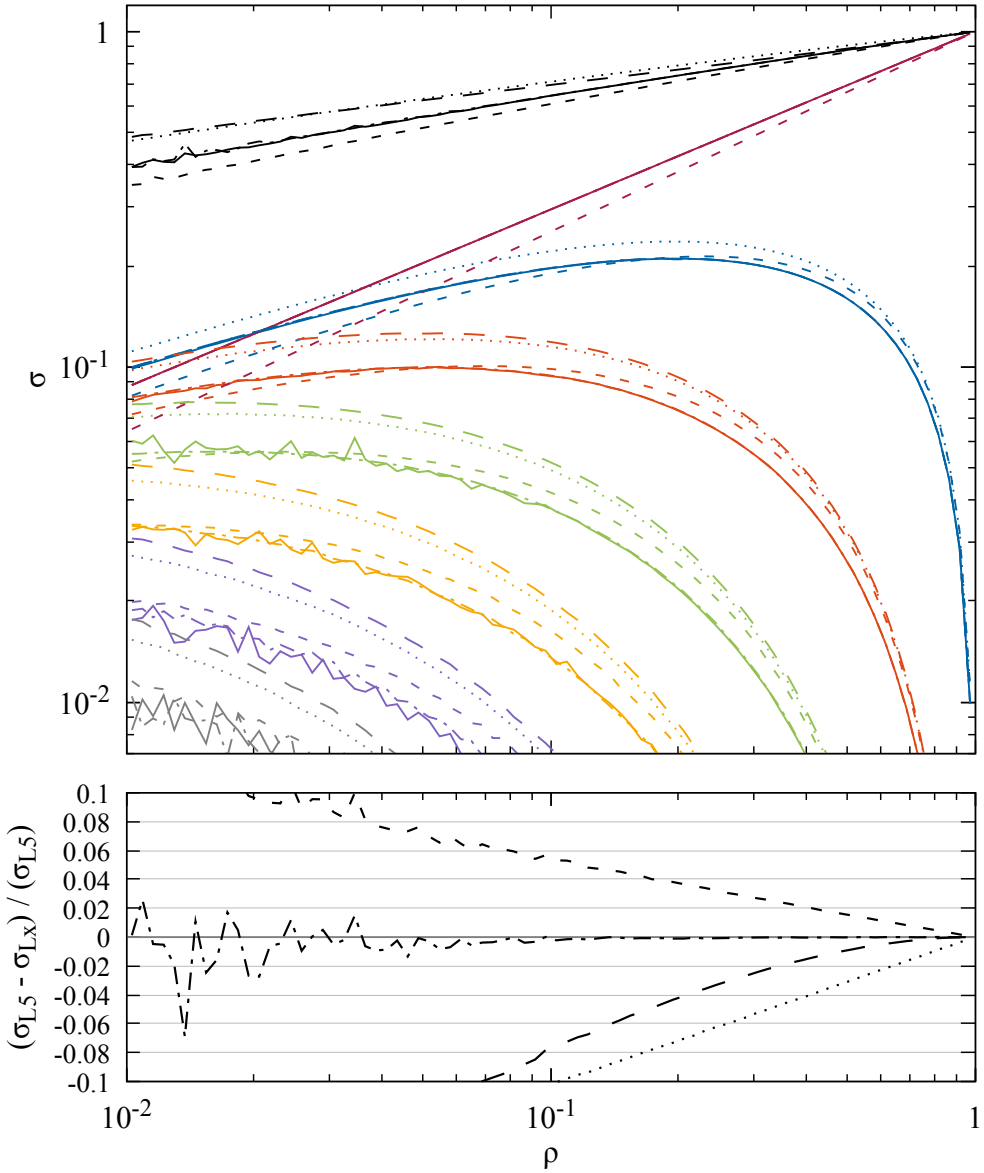
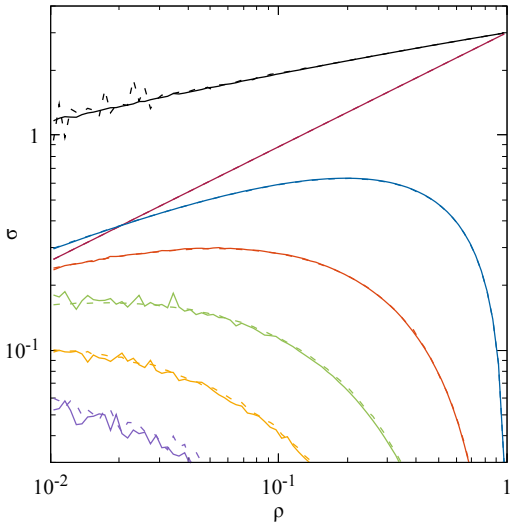
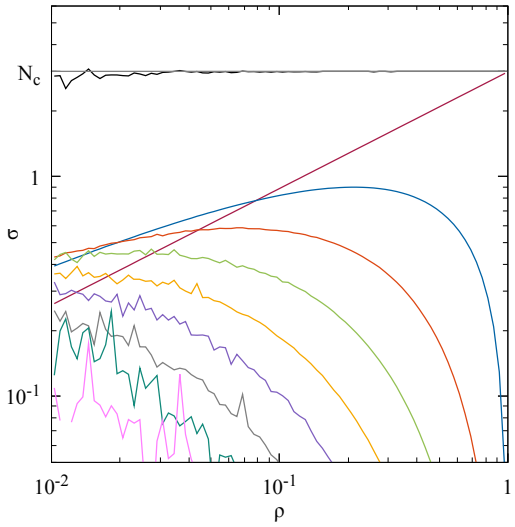


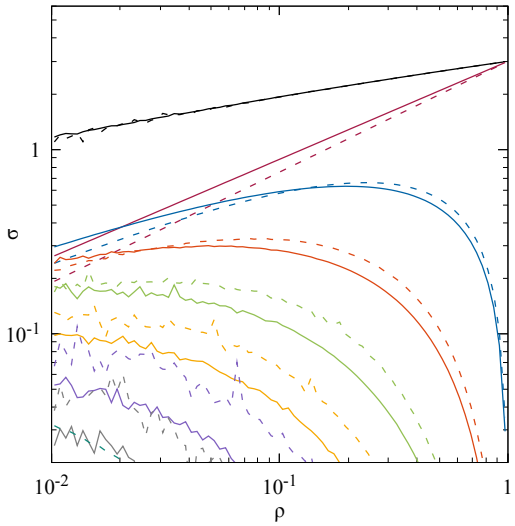
Figure 4.3: The veto cross-section as a function of veto scale for $Z \to q\bar{q}$. Solid: Full colour (L5), Dash-dotted: LC′ + FCR (L4), Long-dashed: LC′ + LCR + singlets (L3), Dotted: LC′ + LCR (L2), Short-dashed: strict LC (L1). The different coloured curves correspond to different multiplicities (0 up to 6 emissions) and the black curves are the total cross-section. The lower residual plot is for the total cross-section.



(a) Testing Lorentz invariance. The solid curves are for the standard kinematic configuration whilst the dashed curves correspond to a boosted configuration where the dijets have an opening angle of $\pi/4$. The interjet region and collinear cutoff are correspondingly boosted.



(b) Testing unitarity. The gap is eliminated so that real emissions are not vetoed and the fixed multiplicity curves sum up to a constant total (N_c).



(c) Testing cutoff independence. The solid curves correspond to $\lambda = 0.01$ and the dashed curves to $\lambda = 0.005$. Though the results are very different per multiplicity, the total is independent of λ.

Figure 4.4: Control plots for $Z \to q\bar{q}$. In all cases, the upper curves are for the total and the lower curves are the contributions from different multiplicities (from 0 to 5 emissions).

plot compares the veto cross-section obtained in the standard configuration with that obtained by boosting the two hard partons such that their opening angle is $\pi/4$. The same boost is applied to the definition of the interjet region and to the collinear cutoff, i.e. $\lambda = 0.01$ for the standard configuration and $\lambda = 0.00146$ for the boosted configuration, such that the ratio $n_i \cdot n_j/\lambda$ is constant. The second plot shows that when we remove the gap the individual multiplicities give contributions that sum up to N_c. This is a highly non-trivial check of the algorithm. As is the third plot, which checks that our results do not depend on the choice of collinear cutoff over the range in ρ that we explore. This is despite the fact that reducing the collinear cutoff enhances the effect of higher multiplicities, e.g. the dashed-grey 5 emission contribution, corresponding to the smaller cutoff, is similar in size to the solid-purple 6 emissions contribution for the larger cutoff.

We now turn to consider colour singlet production of a pair of gluons, e.g. $H \to gg$. In this case the hard process density matrix is

$$H = |10\rangle\langle 10| - \frac{1}{N_c}|10\rangle\langle 01| - \frac{1}{N_c}|01\rangle\langle 10| + \frac{1}{N_c^2}|01\rangle\langle 01|, \tag{4.1}$$

which is represented by the matrix whose elements are $(\mathbf{H})_{ij} = \langle i|H|j\rangle$ where

$$\mathbf{H} = \begin{pmatrix} 1/N_c^2 & -1/N_c \\ -1/N_c & 1 \end{pmatrix} \tag{4.2}$$

and the scalar product matrix is $S = \mathbb{I} = \sum_i |i]\langle i|$, which is represented by the matrix whose elements are $(\mathbf{S})_{ij} = [i|S|j] = \langle i|j\rangle$ where

$$\mathbf{S} = \begin{pmatrix} N_c^2 & N_c \\ N_c & N_c^2 \end{pmatrix}. \tag{4.3}$$

The total cross-section is $\sigma_0 = \mathrm{Tr}(\mathbf{HS}) = N_c^2 - 1$. The veto cross section is presented in Fig. 4.5 and it exhibits the remarkable property first noted in [112], and further verified in [111], whereby the full colour result is correctly described by the leading colour approximation up to a factor of $(N_c^2 - 1)/N_c^2$. We cannot explain this result but it is noteworthy that three of the four contributions to $\mathbf{H}$ do not evolve at all (the three terms involving singlet gluons). In other words, the singlet gluons do not emit and their colour evolution is trivial. However, it is surprising that even the multi-gluon emission cross-section with full

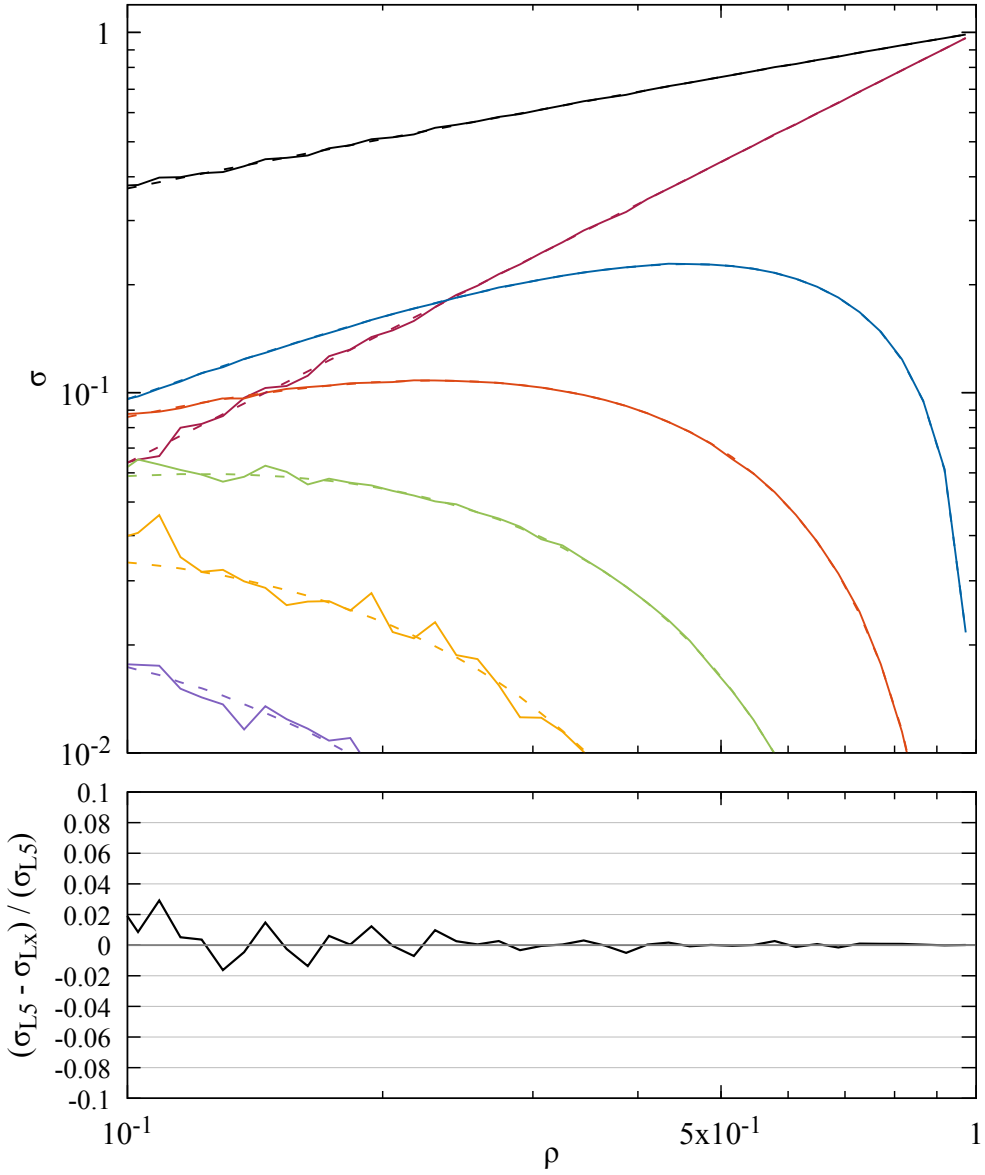


Figure 4.5: The veto cross-section as a function of veto scale for $H \to gg$. Solid: Full colour, Dashed: leading colour. The different coloured curves correspond to different multiplicities (0 up to 5 emissions) and the black curves are the total cross-section. The lower residual plot is for the total cross-section.

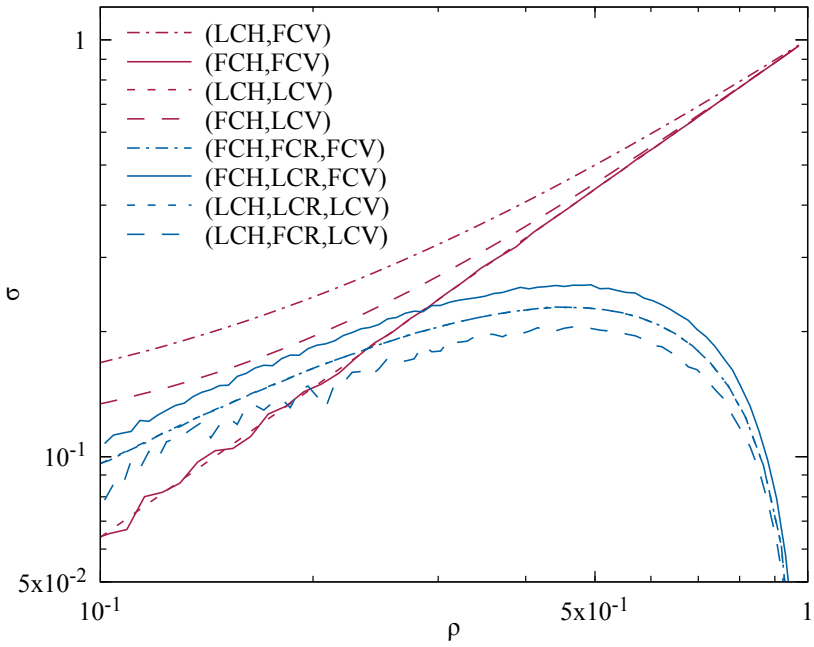
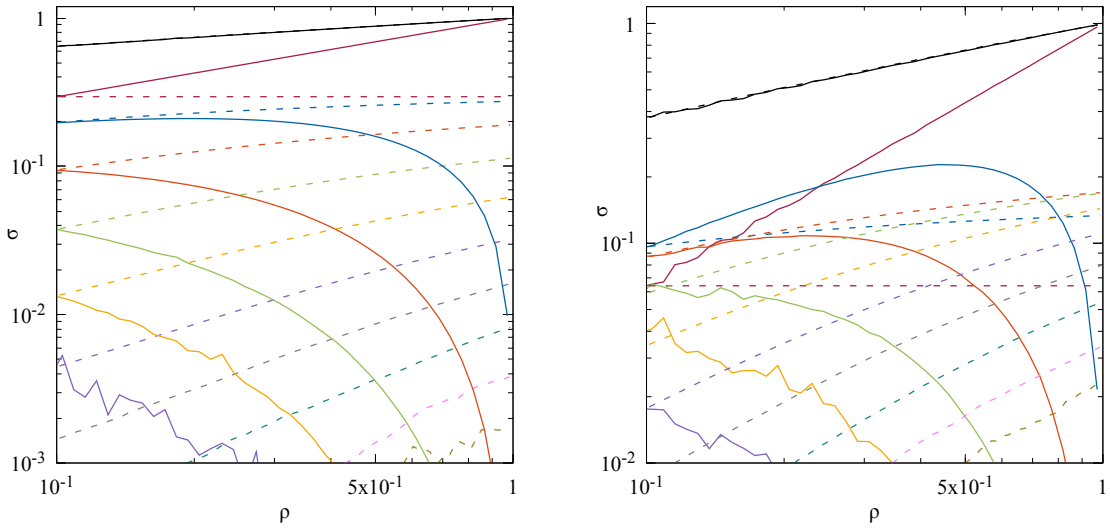


Figure 4.6: The veto cross-section as a function of veto scale for $H \to gg$. The red curves correspond to 0 emissions, and the blue curves to 1 emission.

colour evolution still only differs from the leading colour result by a factor $(N_c^2-1)/N_c^2$. Note that the approximations equivalent to L1, L2 and L3 in the quark case (but with LC rather than LC′) are all equivalent in the gluon case. Approximation L4 differs because of the way that real emissions from rings and strings can move the density matrix away from the leading colour form. As a result, for purely gluonic processes, we shall refer to the LCR (leading colour real) approximation, which excludes rings and strings, and FCR (full colour real). We shall also refer to LCH (leading colour in the hard scatter matrix) and FCH (full colour in the hard scatter matrix), and to LCV (leading colour virtuals) and FCV (full colour virtuals). In this language, Fig. 4.5 illustrates that (FCH,FCR,FCV) is equal to (LCH,LCR,LCV) up to a factor $(N_c^2-1)/N_c^2$. This means there is a conspiracy between the real and virtual evolution that leads to a cancellation of the sub-leading colour.

This intriguing result is explored further in Fig. 4.6, which shows results for zero and one gluon emission. In the zero emission case, the two new curves correspond to (LCH,FCV) evolution (the upper dot-dashed curve) and (FCH,LCV) (long dashed). Both of these are very different from the (FCH,FCV) and (LCH,LCV) curves (solid and short dashed). Remember that we always re-scale the LCH curves to match the total cross-section, i.e. to ensure agreement with FCH at $\rho = 1$. The message for zero emissions is clear: it is all or nothing, in that one must evolve the complete hard scatter matrix with full colour evolution or the leading colour hard scatter matrix with leading colour evolution.



(a) The veto cross section for $Z \to q\bar{q}$.

(b) The veto cross section for $H \to gg$.

Figure 4.7: Comparison of the dedicated (solid curves) and event generator (dashed curves) modes of `CVolver` broken down by multiplicity. The total cross sections agree perfectly whilst those for individual multiplicities are very different.

The way in which the conspiracy works is illustrated by the one gluon emission curves in the figure. We now show the (FCH,LCR,FCV) (solid) and (LCH,FCR,LCV) (long dashed) curves. Remarkably, the (FCH,LCR,FCV) curve is larger than the correct result ((FCH,FCR,FCV)=(LCH,LCR,LCV)) by a factor $N_c^2/(N_c^2-1)$ and the (LCH,FCR,LCV) result is smaller than the correct result by a factor $(N_c^2-1)/N_c^2$. This scaling persists at least up to 5 emissions and indicates that the (very simple) hard process normalisation is somehow encoded in the (very complicated) real emissions.

We conclude this section by showcasing the operation of `CVolver` as an event generator. Fig. 4.7 shows the veto cross-section for $Z \to q\bar{q}$ and $H \to gg$ broken down by multiplicity. Notice that there is perfect agreement between the two modes for the total cross-section, which is built up from a highly non-trivial sum over multiplicities. The dashed curves correspond to the veto cross-sections for a fixed number of emissions and they exhibit the correct physical behaviour, i.e. flat for zero emissions and falling as the veto increases in severity (i.e. as ρ falls) for higher multiplicities. In this case, emissions may be in the angular region where the veto operates provided they have energy less than Q_0. This is in stark contrast to the dedicated mode where all emissions are necessarily emitted at angles outside of the veto region and Q_0 is fixed by the energy of the first emission into the veto region. A detailed study of colour evolution using the dedicated event generator mode will be presented in Chapter 5. For the remainder of this chapter, we shall focus on results obtained using the dedicated mode.

4.2 Two-to-Two processes

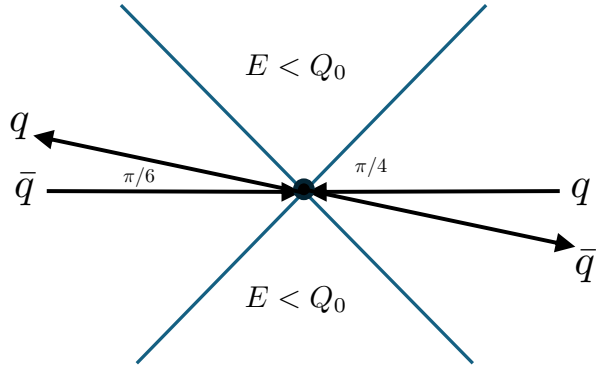


Figure 4.8: Two outgoing partons are produced back-to-back and at an angle of $\pi/6$ relative to the axis defined by the incoming particles. Radiation is vetoed in the central region.

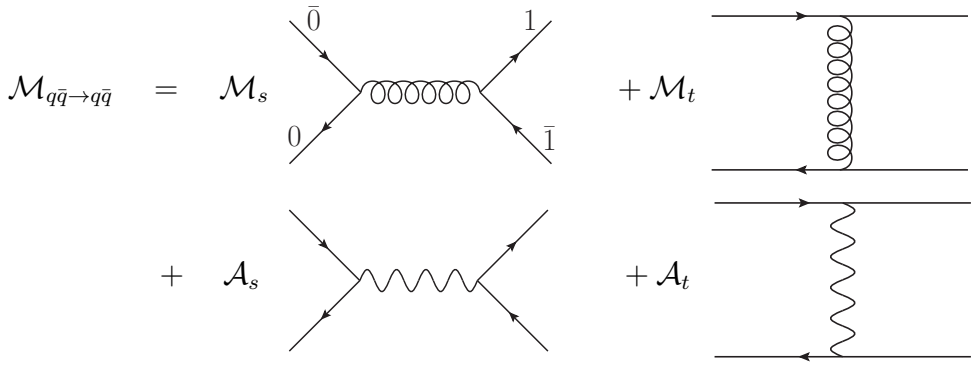
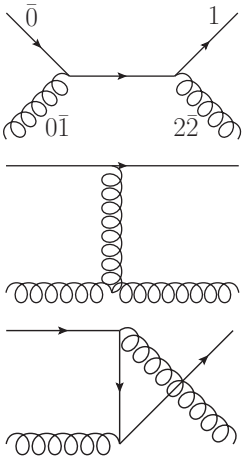


Figure 4.9: The diagrams contributing to the amplitude for $q\bar{q} \to q\bar{q}$.

We now turn our attention to two-to-two hard scattering processes. The goal is to explore the dynamics of sub-leading colour in a controlled and systematic way. To this end we will analyse $q\bar{q} \to q\bar{q}$, $qg \to qg$ and $gg \to gg$. We will consider a veto region defined as in Fig. 4.8. This is the same configuration as that studied in [112] and it allows us to compare with the results in that paper. We will also consider a more asymmetric configuration. We do not include the $i\pi$ terms due to Coulomb gluon exchanges in the virtual evolution operator. This is because the cross-section is divergent in the limit $\lambda \to 0$ in the soft approximation due to the presence of super-leading logarithms [87, 88]. We do see this as a logarithmic dependence upon the collinear cutoff for all two-to-two scattering processes. This dependence cancels in the processes we considered in the previous section, since the Coulomb gluon contribution precisely cancels in those cases. We aim

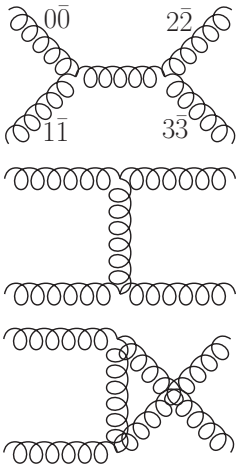


$$= |021\rangle - \frac{1}{N_c}|120\rangle - \frac{1}{N_c}|012\rangle + \frac{1}{N_c^2}|102\rangle$$

$$= |210\rangle - |021\rangle$$

$$= |210\rangle - \frac{1}{N_c}|120\rangle - \frac{1}{N_c}|012\rangle + \frac{1}{N_c^2}|102\rangle$$

(a) $qg \to qg$



$$= |3012\rangle + |1230\rangle - |1302\rangle - |2031\rangle$$

$$= |3201\rangle + |2310\rangle - |1302\rangle - |2031\rangle$$

$$= |3201\rangle + |2310\rangle - |3012\rangle - |1230\rangle$$

(b) $gg \to gg$

Figure 4.10: The diagrams and colour flows for the qg and gg scattering processes.

to study Coulomb exchanges in more detail in a future analysis. In this section, we focus on presenting our results without detailing the extensive cross-checks performed to verify their validity. Nevertheless, considerable effort has gone into ensuring their correctness, and selected examples of these cross-checks are provided in Appendix A.

4.2.1 $q\bar{q} \to q\bar{q}$

The matrix element for $q\bar{q} \to q\bar{q}$ is illustrated in Fig. 4.9 and, in the colour flow basis, it can be written

$$\mathcal{M}_{q\bar{q}\to q\bar{q}} = \left(\frac{\mathcal{M}_s}{2} - \frac{\mathcal{M}_t}{2N_\mathrm{c}} + \mathcal{A}_t\right)|10\rangle + \left(\frac{\mathcal{M}_t}{2} - \frac{\mathcal{M}_s}{2N_\mathrm{c}} + \mathcal{A}_s\right)|01\rangle. \tag{4.4}$$

We shall not concern ourselves with the kinematic factors ($\mathcal{M}_{s,t}$ and $\mathcal{A}_{s,t}$), preferring instead to focus on the contributions from specific terms in the density matrix. The hard scatter matrix includes contributions from $|01\rangle\langle 01|$, $|10\rangle\langle 10|$ and $|01\rangle\langle 10|$ and we shall now investigate the veto cross-section arising from each of these three possibilities.

In Fig. 4.11 we show the L1 to L5 curves for the case $H = |01\rangle\langle 01|$. This is the configuration where the outgoing particles are colour connected and the incoming particles are colour connected. At leading colour in the hard process, this corresponds to t-channel gluon exchange (or s-channel γ/Z exchange). The lower pane shows the residuals for the total rates and indicates that leading colour LC′ (L2) agrees best with the full colour result. This is in contrast to the $Z \to q\bar{q}$ case, where L4 provided the best approximation. Notice that the agreement between L1 and L2 is only for the total. Even at zero emissions there is a big difference between red L5 and red L4=L3=L2, which implies large effects from virtual swaps. The differences between L5 and L4 show that virtual swaps are also involved in the cancellations beyond L2.

In Fig. 4.12 we show the L1 to L5 curves for the case $H = |10\rangle\langle 10|$. This is the configuration where the incoming quark is colour connected to the outgoing quark and similarly for the incoming and outgoing antiquarks. At leading colour in the hard process, this corresponds to s-channel gluon exchange (or t-channel γ/Z exchange). In contrast to the previous case, it is L4 that agrees best with the full colour result and the leading colour LC′ approximation fails. This conclusion appears to be in agreement with [112] who study $qq \to qq$ with t-channel electroweak boson exchange and find that the mean field approximation agrees best with full colour[1]. Contrary to Fig. 4.11, this time the agreement L4=L5 works at the level of each multiplicity, which shows there is little effect from virtual swaps, i.e. it seems the evolution does not favour swapping to the other colour configuration.

We now turn to the interference contribution arising from $H = |10\rangle\langle 01|$, which is

[1]Though [112] say that their result may be unreliable due to lattice artifacts.

illustrated in Fig. 4.13. Only L3, L4 and L5 have any emissions. That is because the L1 and L2 approximations require that emission be off dipoles and in this case there are no dipoles to emit from. Hence the L1 and L2 total cross-section is equal to the zero emission cross-section. The L3, one-emission curve is the blue, long-dashed curve towards the bottom of the plot. It is negative and we plot the absolute value. As before, the zero emission curves for L2, L3 and L4 are all coincident. Clearly, none of the LC′ approximations is good with differences in all cases exceeding 50% at $\rho = 0.1$.

We can also explore the veto cross-section for s-channel and t-channel gluon exchanges, which is achieved by combining the previous results. The corresponding hard scatter matrices are proportional to

$$\mathbf{H}_s = \begin{pmatrix} 1/N_c^2 & -1/N_c \\ -1/N_c & 1 \end{pmatrix} \text{ and } \mathbf{H}_t = \begin{pmatrix} 1 & -1/N_c \\ -1/N_c & 1/N_c^2 \end{pmatrix}. \tag{4.5}$$

The s-channel contribution is the complete lowest-order QCD contribution to $q\bar{q} \to Q\bar{Q}$. The corresponding veto cross-section is shown in Fig. 4.14. Again we see that, though the approximations do steadily improve, none agrees well with the full colour result. In fact in all cases, the difference exceeds 10% at $\rho = 0.1$. The equivalent curves in the case of t-channel gluon exchange are in Fig. 4.15. This would be the complete lowest-order QCD contribution to $q\bar{Q} \to q\bar{Q}$. Again none of the L1–L3 approximations is good and the L4 approximation is breaking down at low ρ.

The results we just presented for s and t-channel gluon exchange evolved from the full-colour hard-scatter matrices in Eq. (4.5). Starting instead from the leading-colour hard-scatter matrices for the L1–L4 approximations (i.e. we still use the full-colour hard-scatter matrix for L5) we find that the L1–L4 approximations fare significantly better than before[2]. In fact, the t-channel contribution shown in Fig. 4.16 again confirms a result first found in [112] that the full-colour evolution of the full-colour hard process is very well approximated by the leading-colour evolution of the leading-colour hard process (i.e. the strictly leading colour approximation). This is the same result we found for $H \to gg$, though in this case is perhaps even more surprising since the agreement is found by taking $C_F \to N_c/2$ even in the diagonal part of the virtual evolution. Indeed, this change shifts the slope of the zero emission curve in such a way to compensate shifts in the higher multiplicity contributions, which is in contrast to Fig. 4.6 where the agreement occurs for each multiplicity. The success of the strictly leading colour approximation does not hold

[2]The L1–L4 curves on Fig. 4.16 are identical to those on Fig. 4.11 and those on Fig. 4.17 are identical to those on Fig. 4.12. Only the solid (L5) curves differ.

for s-channel gluon exchange, as illustrated in Fig. 4.17.

Let us now change the kinematics and put the final state quark at rapidity $y = 4$ and the final state antiquark at $y = -2$. The veto region is the interjet region $-1.3 < y < 3.3$. This asymmetric, low-angle scenario leads to very different results from the previous configuration. The $|10\rangle\langle 10|$ configuration is much less likely to radiate into the veto region than the $|01\rangle\langle 01|$[3] and this immediately renders all of the LC′ approximations inadequate whenever the $|01\rangle\langle 01|$ contribution is important since none are able to flip the colour of the $|01\rangle\langle 01|$ configuration, via a virtual exchange, into the $|10\rangle\langle 10|$ state. We expect that this effect would be even more dramatic had we included Coulomb exchanges. Note that in a hadron-hadron collision, the $|01\rangle\langle 01|$ contribution would be leading since t-channel gluon exchange dominates for low-angle scattering. Conversely, the L4 approximation is very good for the evolution of the $|10\rangle\langle 10|$ configuration. These features are all illustrated in Fig. 4.18–Fig. 4.20. Fig. 4.19 also shows that the L1 (strictly leading colour) approximation is a good approximation for the evolution of the $|10\rangle\langle 10|$ configuration.

For the evolution of the s-channel gluon exchange diagram starting from the full-colour hard scatter matrix (H_s), we find that all of the LC′ approximations are at least 20% above the full colour result at $\rho = 0.1$. For evolution of the dominant t-channel gluon exchange diagram, the situation is, as expected, even more extreme and none is correct within a factor of 2 at this value of ρ. Starting from the leading colour hard-scatter again improves matters and the L1 strictly leading colour curve once again agrees with the full colour result. These features are illustrated in Fig. 4.21–Fig. 4.24.

A word on the control that we have over these results. We have checked that the d-approximation is convergent (the $d = 2$ approximation is sufficient for all of the plots presented so far) and we have checked that our results agree perfectly with independent analytic calculations for zero emissions and one emission. That the $d = 2$ approximation is sufficient for the full-colour curves is at first a surprise since $d = 4$ is formally required at this order in $1/N_c$. However, the lower multiplicities have the largest relative contribution from larger d for combinatoric reasons, i.e. there are many more possible $d = 1$ contributions at high multiplicities, and for these we can check explicitly that $d = 2$ is good. We have also checked that our results are precisely independent of the collinear cutoff. For the symmetric scenario we find that $\lambda = 0.01$ is sufficient and for the asymmetric scenario we need to take a smaller value, $\lambda = 0.001$.

[3] To a lesser extent this was also true for the previous kinematics.

Channel	FCH		LCH	
	symmetric	asymmetric	symmetric	asymmetric
s	L4:$\sim$15%	(L1,L2,L3):$\sim$35%	L2:$\sim$10%	(L1,L2,L3):$\sim$50%
t	$\approx$L4	=L4	=L1 [112]	=L1, L2, L3
u	L4:$\sim$15%	(L1,L2,L3):$\sim$35%	L2:$\sim$15%	(L1,L2,L3):$\sim$50%
st	$\approx$L4	$\approx$L4	$\approx$L1	$\approx$L1, L2, L3
su	L1:$\sim$130%	(L1,L2,L3):$\sim$80%	N/A	N/A
tu	$\approx$L4	$\approx$L4	$\approx$L1, L3	$\approx$L1, L2, L3

Table 4.1: Summary of results for $qg \to qg$. In all cases, the comparison is with the full-colour evolution starting from the full-colour hard scatter matrix modulo an overall normalization. In the cases were none of the approximations successfully reproduce the full colour result, the best performing approximations are listed along their deviation from the full colour result at $\rho = 0.1$.

4.2.2 $qg \to qg$

Now we turn our attention to quark-gluon scattering. There are 3 diagrams to consider, as illustrated in Fig. 4.10a and the corresponding hard scatter matrices are now rank 6. In total there are 6 distinct contributions corresponding to the s, t and u-channel contributions plus the three (st, su and tu) interference contributions. In what follows we will be systematic and explore the contributions to the veto cross-section for each of these. We will do this starting from full-colour in the hard process (FCH) and also leading-colour in the hard processes (LCH). And also for the symmetric and asymmetric kinematic configurations that we explored in the previous section. In total this means we will present 24 plots in Fig. 4.25 to Fig. 4.46. Table 4.1 summarizes the key features of all 24 plots.

In the asymmetric configuration, we find that L1 $\approx$ L2 $\approx$ L3. This is because in this configuration the quark-quark pair is the closest in angle, suppressing that dipole relative to the others, and reducing the impact of singlet exchanges and emissions.

In both kinematic configurations, starting from the LCH, L1 does very well at approximating the L5 result for the t-, st-, and tu- channels. We also see L4 is within 2% of the full colour result for the same channels starting from the FCH. For the other channels, none of the approximations work. In the symmetric configuration, the approximation L2 is never satisfactory, while in the asymmetric configuration the approximations L1–L3 always perform similarly for the reasons explained before. We should also note that, for the channels where subleading colour effects cannot be described without L5 evolution, we see substantial subleading colour corrections: of order 10–20% for the symmetric case and 30–50% for the asymmetric case at $\rho = 0.1$, when compared to the best performing approximation, for the s- and u- channels. The su- interference has enormous effects and

	θ	ϕ
q	0	0
$\bar{q}$	$2\pi/3$	$5\pi/12$
q	$11\pi/12$	$\pi/2$
$\bar{q}$	$3\pi/4$	$\pi/4$

Table 4.2: The directions of the quarks and anti-quarks produced in $ZZ \to q\bar{q}q\bar{q}$. The veto cross-section is defined by vetoing radiation with $E > Q_0$ outside of cones centred on each parton and with an opening angle $\pi/4$.

needs full colour evolution.

4.2.3 $gg \to gg$

To complete our analysis of two-to-two scattering we study $gg \to gg$. In this case, we compare the full-colour result (FCR,FCV,FCH) with the strictly leading-colour approximation (LCR,LCV,LCH). Since the results do not depend much on whether we consider the symmetric or asymmetric veto cross-section we show results only for the asymmetric case in Fig. 4.47 to Fig. 4.52. The remarkable agreement between strict leading colour and full colour for t-channel processes persists and holds even for the tu and st interference contributions. Given the complexity of the subleading colour effects, the success of the vastly simpler leading colour approximation in the t-channel gluon exchange processes is quite remarkable, especially since we do observe very large differences in the evolution between $d = 0$, $d = 1$ and $d = 2$ and emissions from rings and strings, which are absent in the leading colour approximation, greatly increase the possible trajectories in colour space. It seems there is a widespread and very non-trivial cancellation of sub-leading colour effects for these observables (subject to an overall normalization) the structure of which we will be investigating in more detail in the next chapter.

4.3 Colour singlet production of four jets

For our final study we will look at the production of four jets, as might be produced in $ZZ \to q\bar{q}q\bar{q}$. In this case we fix the kinematics as in Table 4.2.

There are two possible colour flows, corresponding to the two possible production amplitudes: $|01\rangle$ and $|10\rangle$. In the former, the first two particles in Table 4.2 are colour connected and so are the final two. In the latter colour flow, it is the first and last particles (and the second and third particles) in the table that are colour connected. Fig. 4.53 shows

that the L1 and L4 approximations are closest to the full colour result for the evolution of the $|01\rangle\langle 01|$ hard process but both differ by 4% at $\rho = 0.1$. For evolution starting from the crossed-channel $|10\rangle\langle 10|$ contribution, the L4 approximation is within 1% of the full colour result (Fig. 4.54). Finally, in Fig. 4.55 we once again see that none of the approximations is adequate and that approximations L1–L3 differ by 80% from the full colour result.

There is a systematic hierarchy between L1, L2 and L3 present in all processes with quarks. L3 is always closer to L1 than L2 is, which implies that adding real singlet emissions counterbalances part of the effect of LC$'$. Therefore, in every diagram where L1 performs well then L3 will perform better than L2. The only cases where L2 performs better than L3 are those where L1 is not good and the overcorrection of LC$'$ is in the right direction: Fig. 4.11, where L2 performs best, and some cases where none of the approximations work well, specifically Figs. 4.31 and 4.33 and the evolutions of individiual interference terms, Figs. 4.13, 4.20 and 4.55.

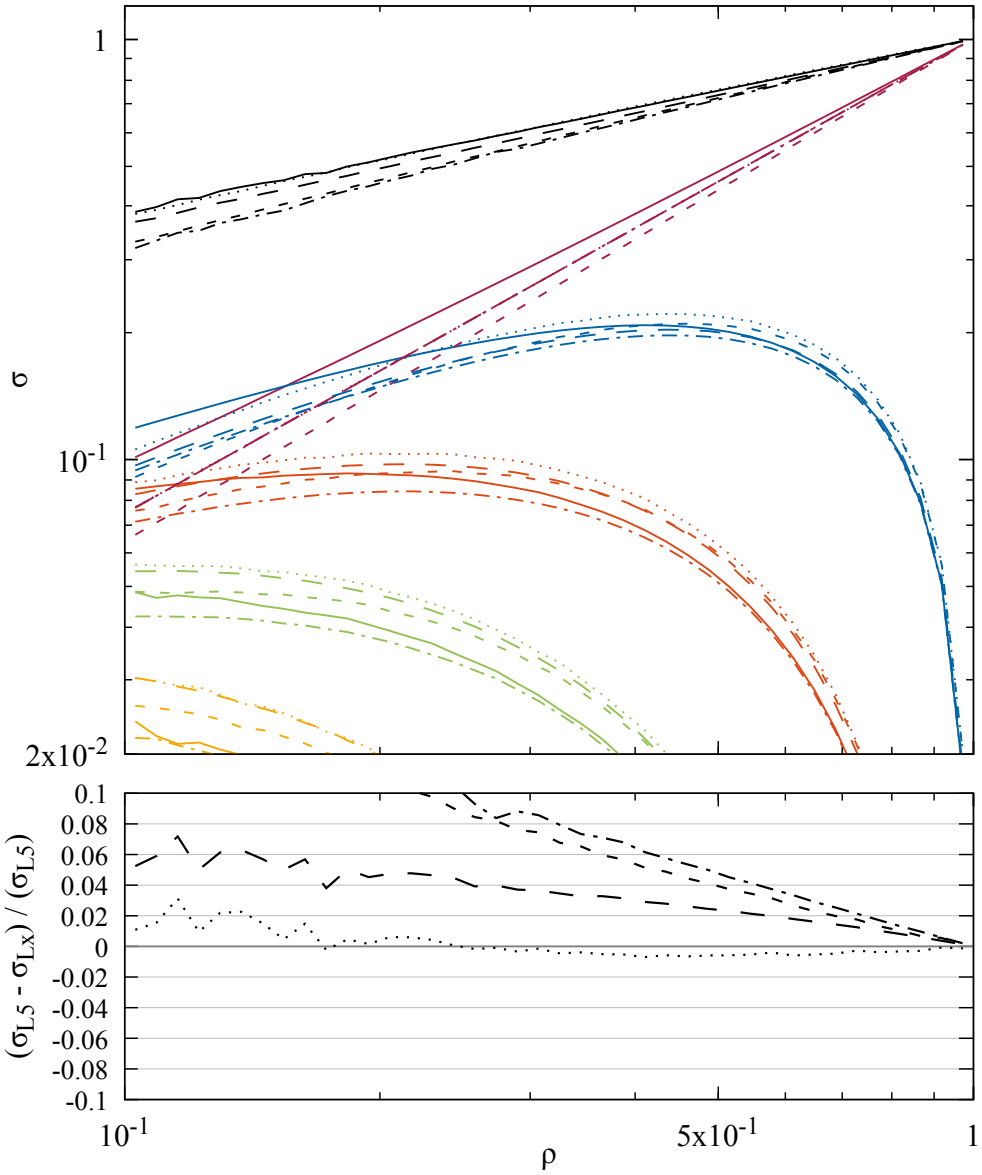


Figure 4.11: The veto cross-section for the $|01\rangle\langle 01|$ contribution to $q\bar{q} \to q\bar{q}$. Solid: Full colour (L5), Dash-dotted: LC′ + FCR (L4), Long-dashed: LC′ + LCR + singlets (L3), Dotted: LC′ + LCR (L2), Short-dashed: strict LC (L1).

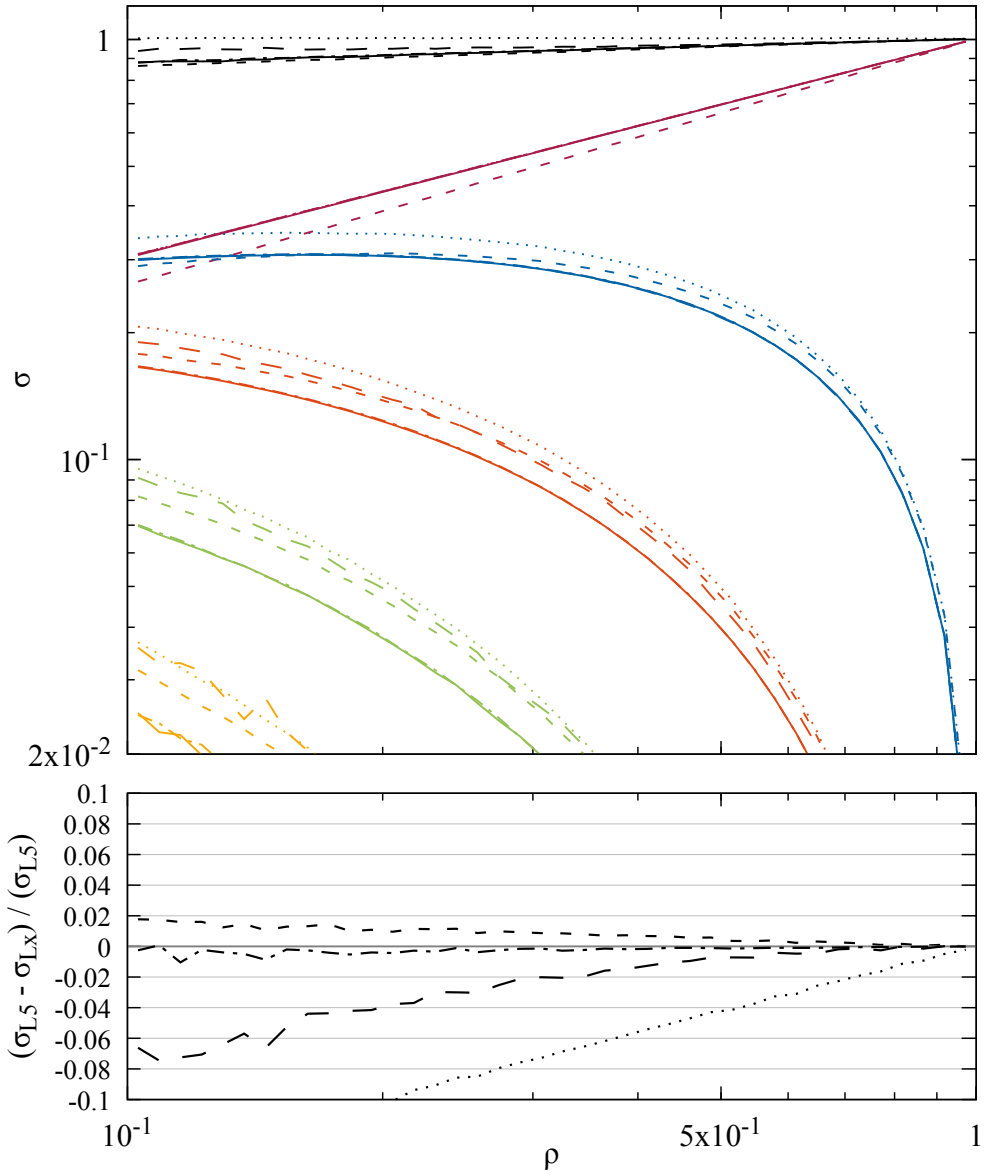


Figure 4.12: The veto cross-section for the $|10\rangle\langle 10|$ contribution to $q\bar{q} \to q\bar{q}$. Solid: Full colour (L5), Dash-dotted: LC′ + FCR (L4), Long-dashed: LC′ + LCR + singlets (L3), Dotted: LC′ + LCR (L2), Short-dashed: strict LC (L1).

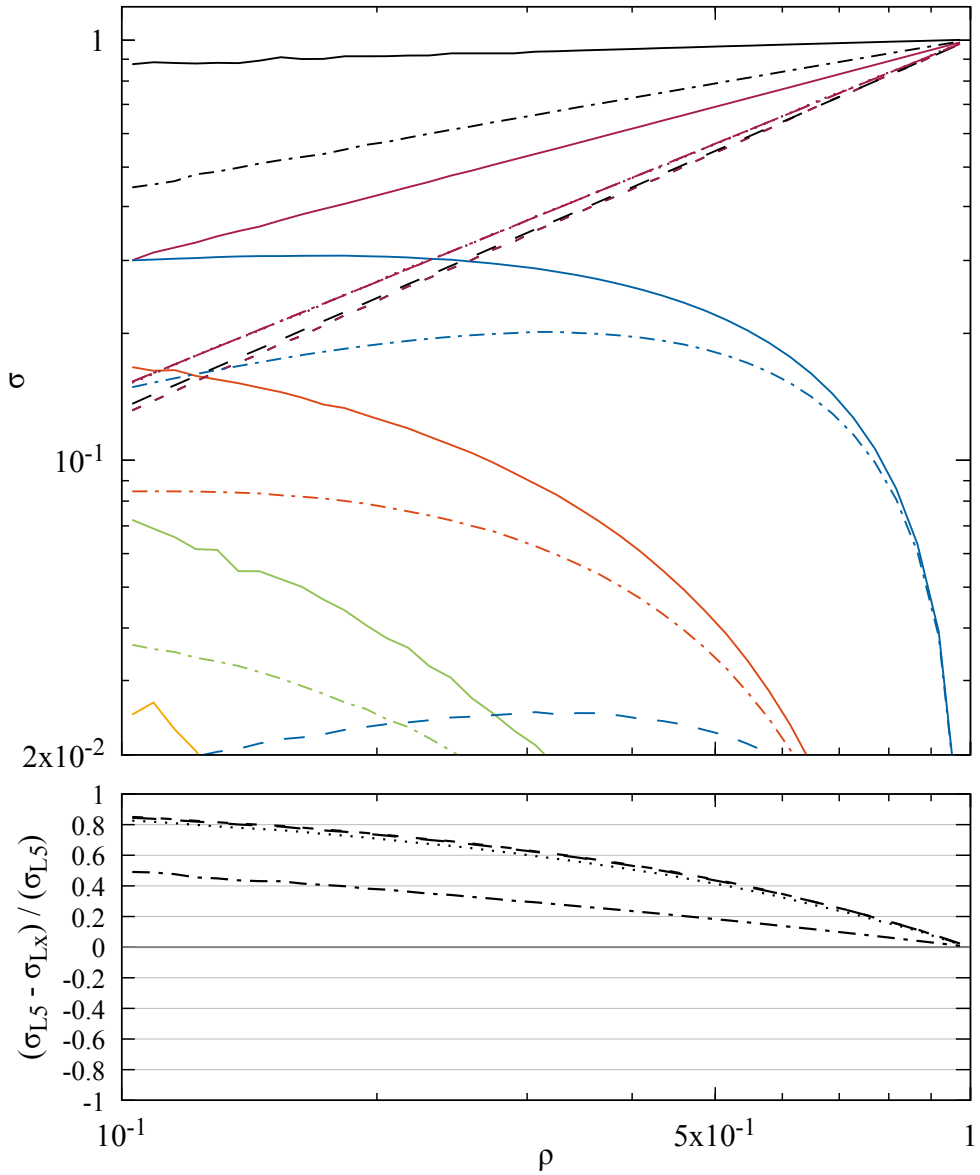


Figure 4.13: The veto cross-section for the $|10\rangle\langle 01|$ (interference) contribution to $q\bar{q} \to q\bar{q}$. Solid: Full colour (L5), Dash-dotted: LC′ + FCR (L4), Long-dashed: LC′ + LCR + singlets (L3), Dotted: LC′ + LCR (L2), Short-dashed: strict LC (L1). The L3, 1 emission cuve (blue, long-dashed) is negative and the absolute value is plotted.

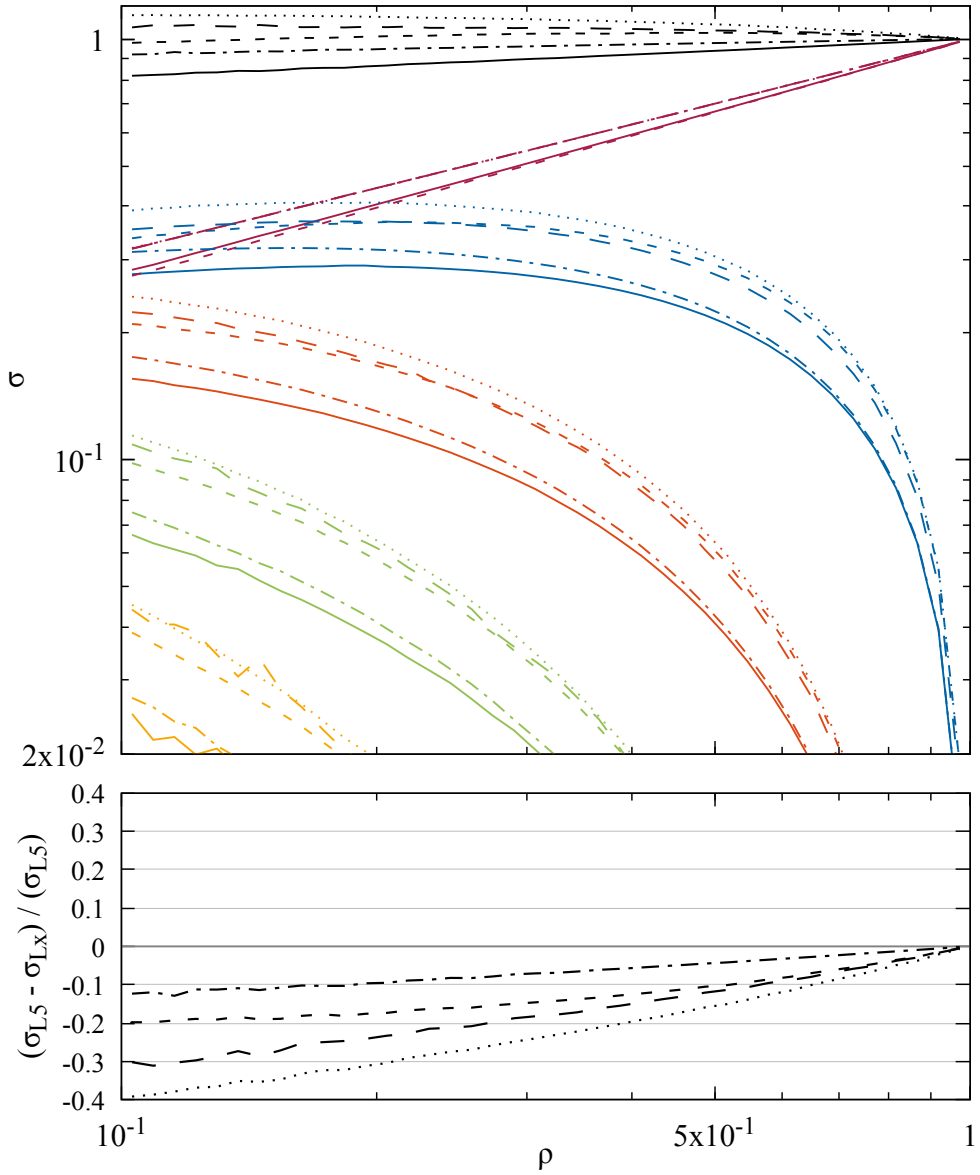


Figure 4.14: The veto cross-section for the s-channel gluon exchange contribution to $q\bar{q} \to q\bar{q}$. Solid: Full colour (L5), Dash-dotted: LC′ + FCR (L4), Long-dashed: LC′ + LCR + singlets (L3), Dotted: LC′ + LCR (L2), Short-dashed: strict LC (L1).

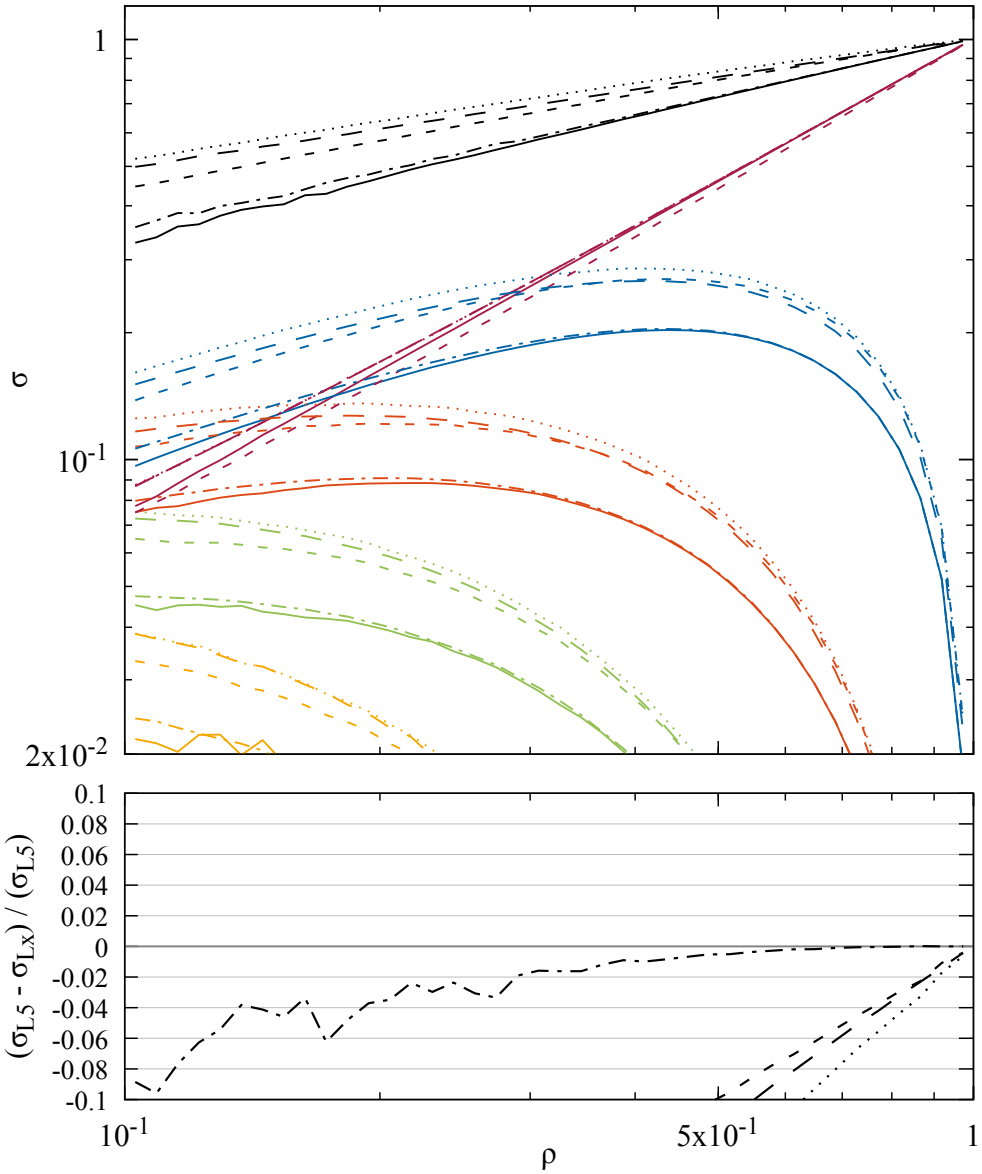


Figure 4.15: The veto cross-section for the t-channel gluon exchange contribution to $q\bar{q} \to q\bar{q}$. Solid: Full colour (L5), Dash-dotted: LC′ + FCR (L4), Long-dashed: LC′ + LCR + singlets (L3), Dotted: LC′ + LCR (L2), Short-dashed: strict LC (L1).

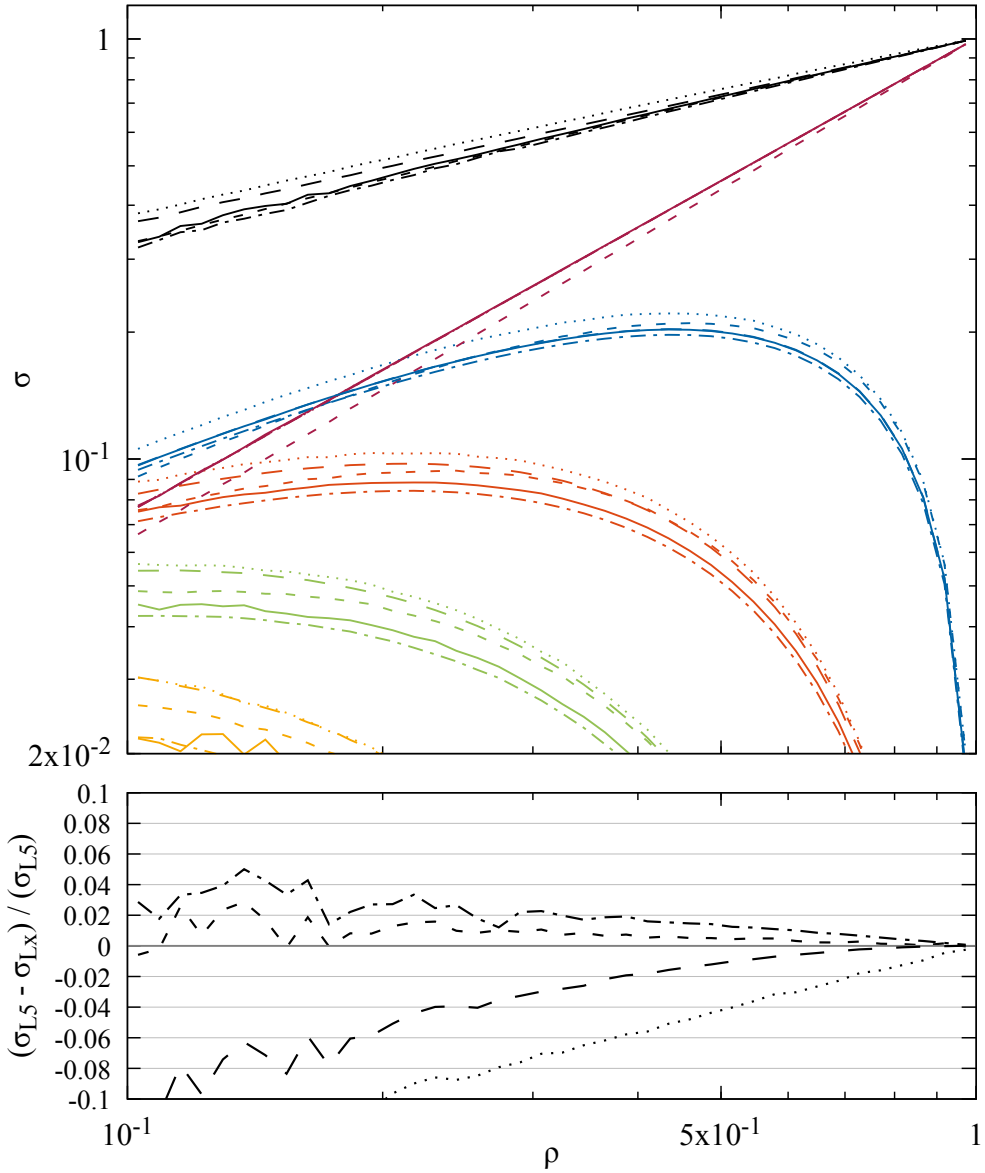


Figure 4.16: The veto cross-section for the t-channel gluon exchange contribution to $q\bar{q} \to q\bar{q}$. Solid: Full colour (L5), Dash-dotted: LC′ + FCR (L4), Long-dashed: LC′ + LCR + singlets (L3), Dotted: LC′ + LCR (L2), Short-dashed: strict LC (L1). For the L1–L4 curves we start the evolution using the leading-colour approximation to the hard-scatter matrix.

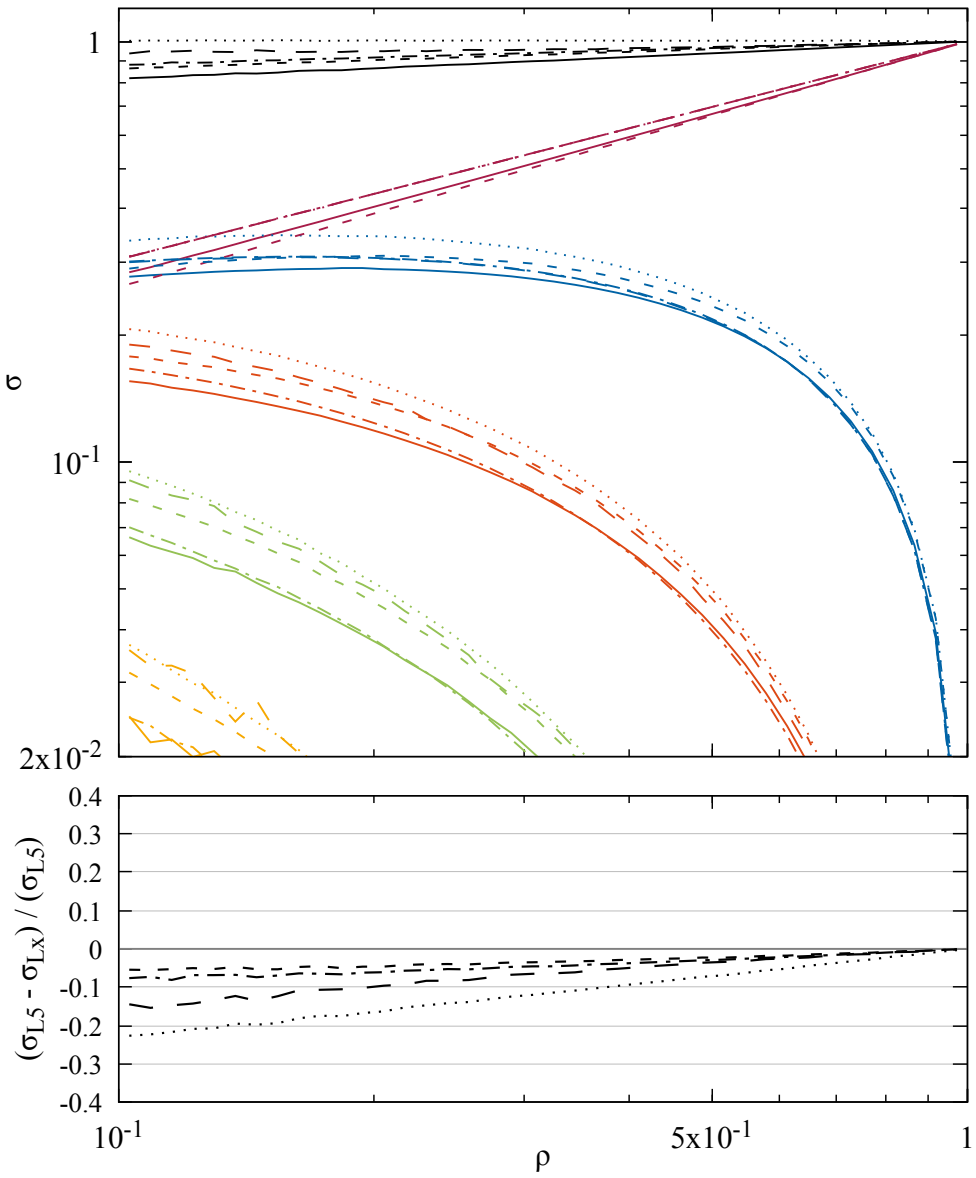


Figure 4.17: The veto cross-section for the s-channel gluon exchange contribution to $q\bar{q} \to q\bar{q}$. Solid: Full colour (L5), Dash-dotted: LC′ + FCR (L4), Long-dashed: LC′ + LCR + singlets (L3), Dotted: LC′ + LCR (L2), Short-dashed: strict LC (L1). For the L1–L4 curves we start the evolution using the leading-colour approximation to the hard-scatter matrix.

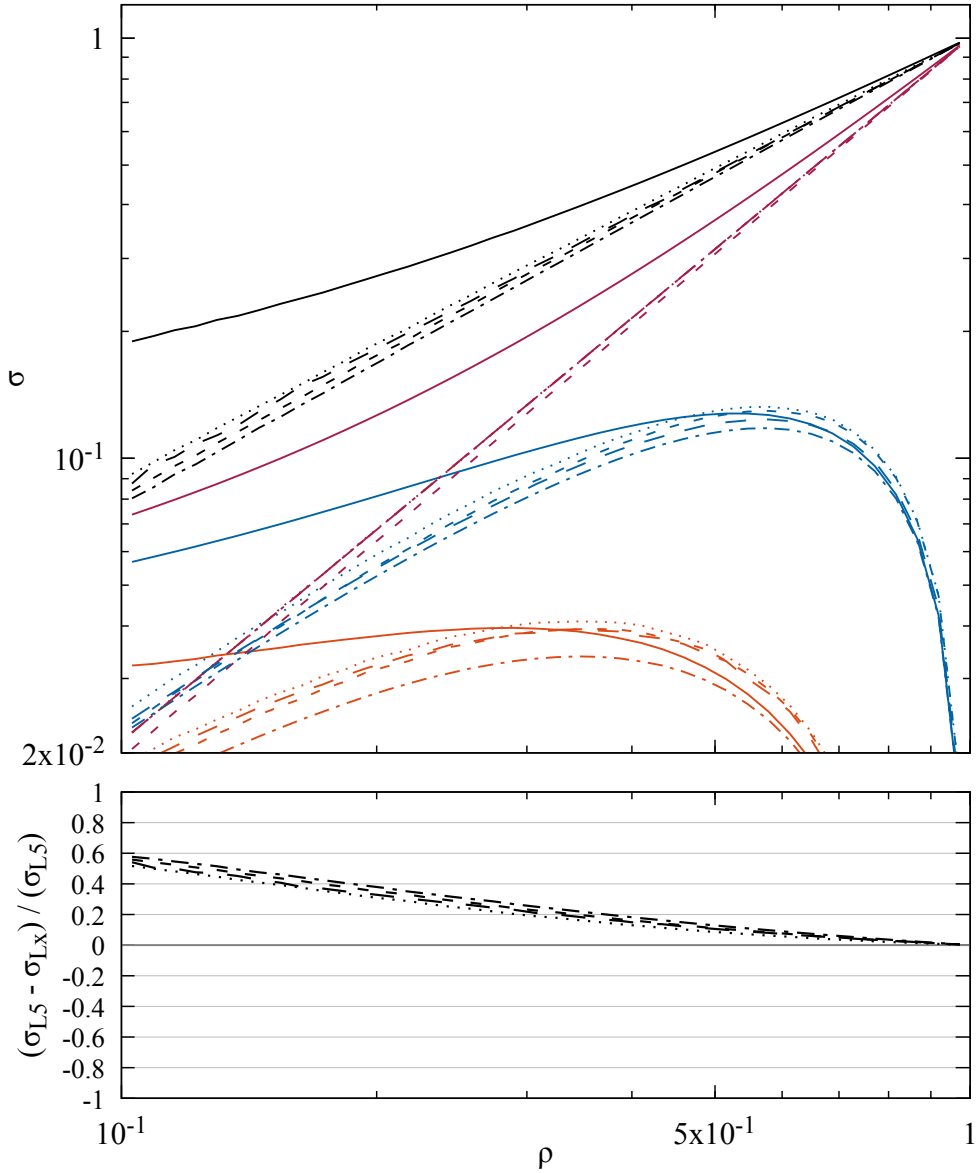


Figure 4.18: The veto cross-section for the $|01\rangle\langle 01|$ contribution to $q\bar{q} \to q\bar{q}$ in the asymmetric configuration. Solid: Full colour (L5), Dash-dotted: LC′ + FCR (L4), Long-dashed: LC′ + LCR + singlets (L3), Dotted: LC′ + LCR (L2), Short-dashed: strict LC (L1).

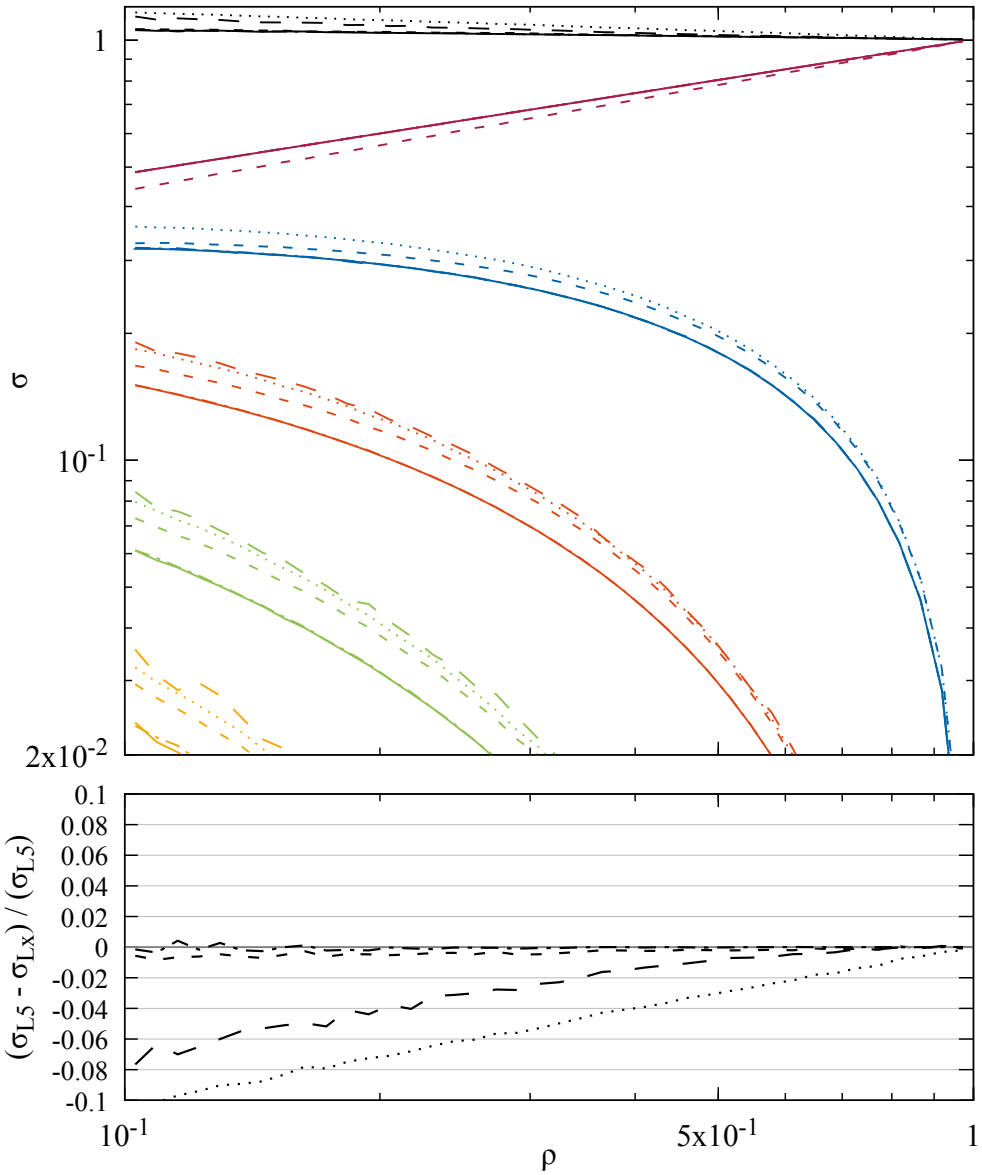


Figure 4.19: The veto cross-section for the $|10\rangle\langle 10|$ contribution to $q\bar{q} \to q\bar{q}$ in the asymmetric configuration. Solid: Full colour (L5), Dash-dotted: LC′ + FCR (L4), Long-dashed: LC′ + LCR + singlets (L3), Dotted: LC′ + LCR (L2), Short-dashed: strict LC (L1).

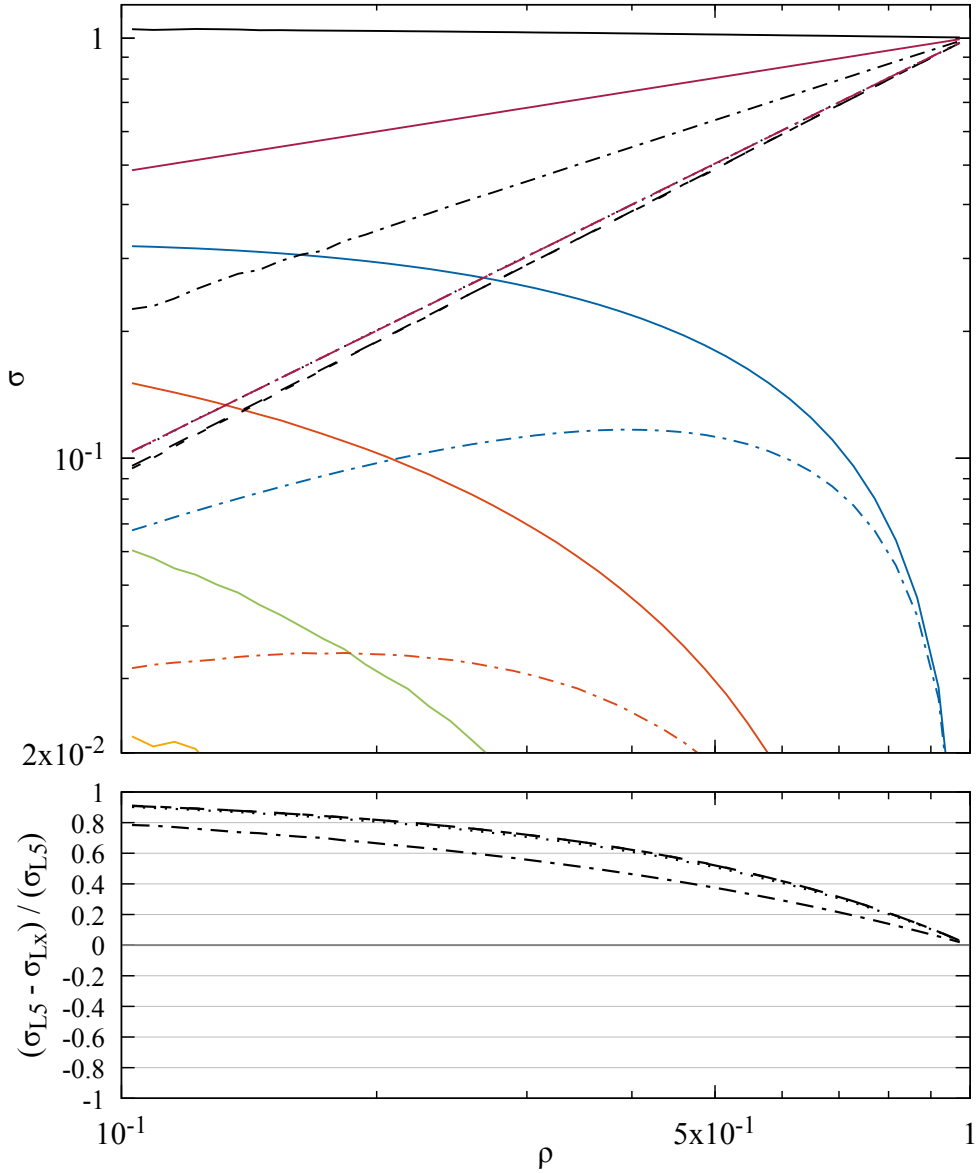


Figure 4.20: The veto cross-section for the $|10\rangle\langle 01|$ (interference) contribution to $q\bar{q} \to q\bar{q}$ in the asymmetric configuration. Solid: Full colour (L5), Dash-dotted: LC′ + FCR (L4), Long-dashed: LC′ + LCR + singlets (L3), Dotted: LC′ + LCR (L2), Short-dashed: strict LC (L1).

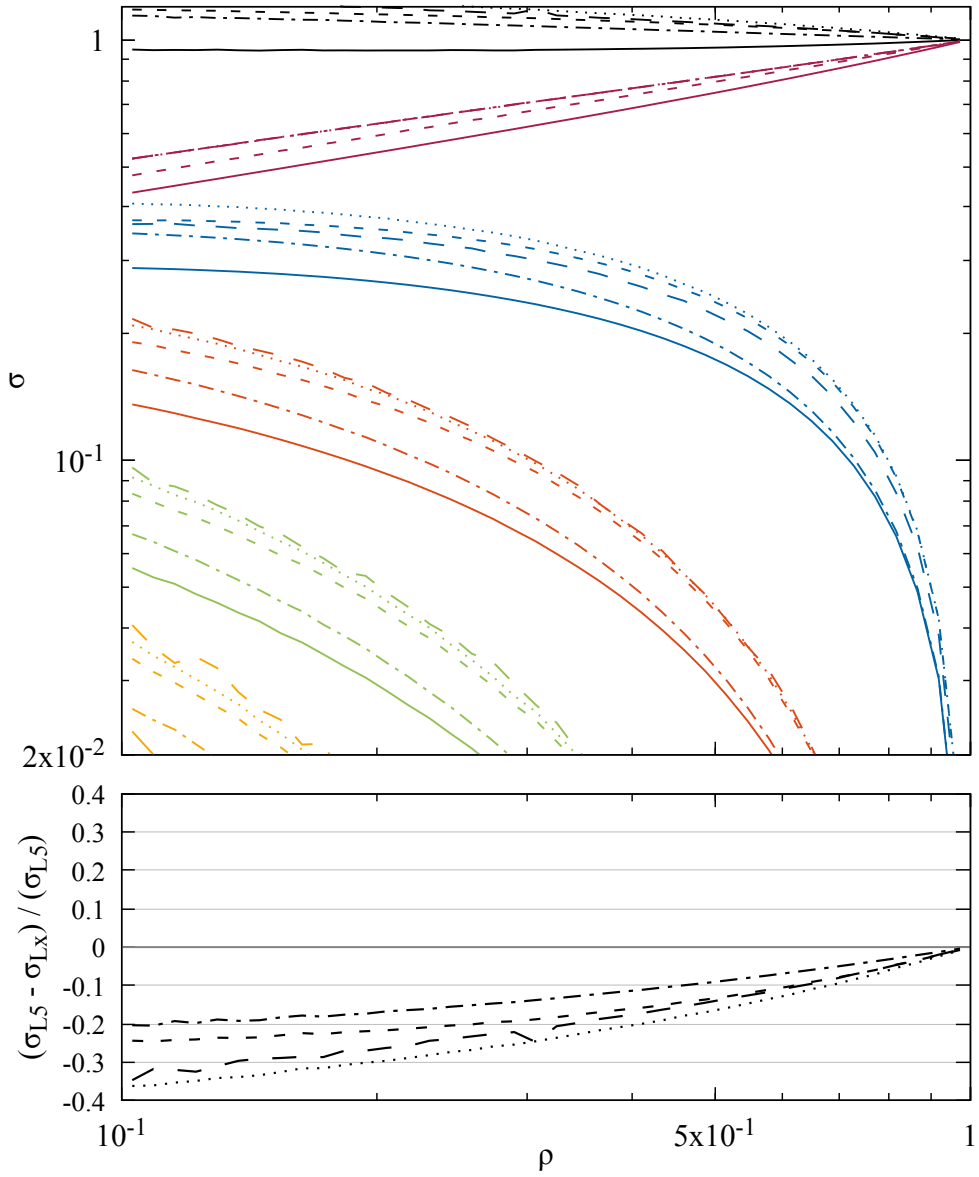


Figure 4.21: The veto cross-section for the s-channel gluon exchange contribution to $q\bar{q} \to q\bar{q}$ in the asymmetric configuration. Solid: Full colour (L5), Dash-dotted: LC′ + FCR (L4), Long-dashed: LC′ + LCR + singlets (L3), Dotted: LC′ + LCR (L2), Short-dashed: strict LC (L1).

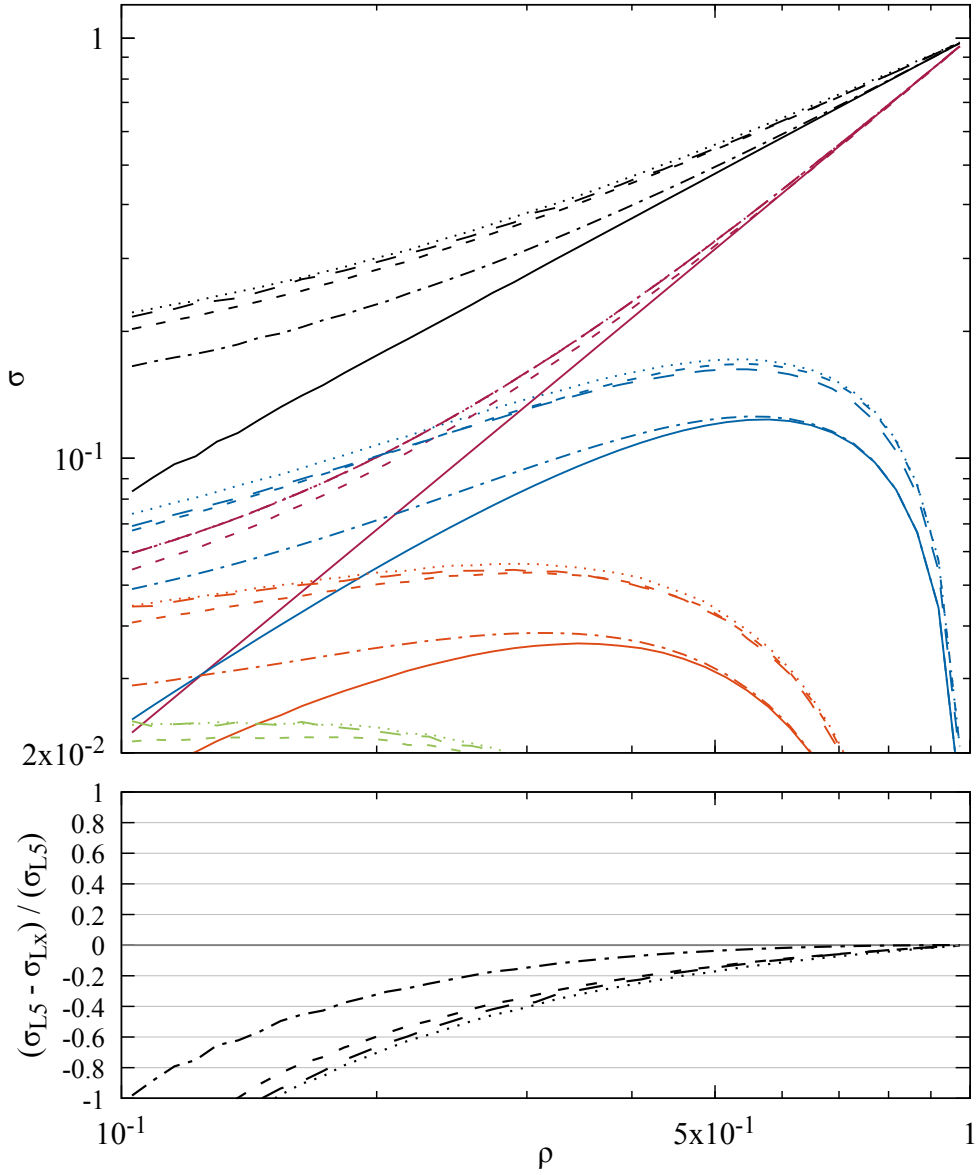


Figure 4.22: The veto cross-section for the t-channel gluon exchange contribution to $q\bar{q} \to q\bar{q}$ in the asymmetric configuration. Solid: Full colour (L5), Dash-dotted: LC′ + FCR (L4), Long-dashed: LC′ + LCR + singlets (L3), Dotted: LC′ + LCR (L2), Short-dashed: strict LC (L1).

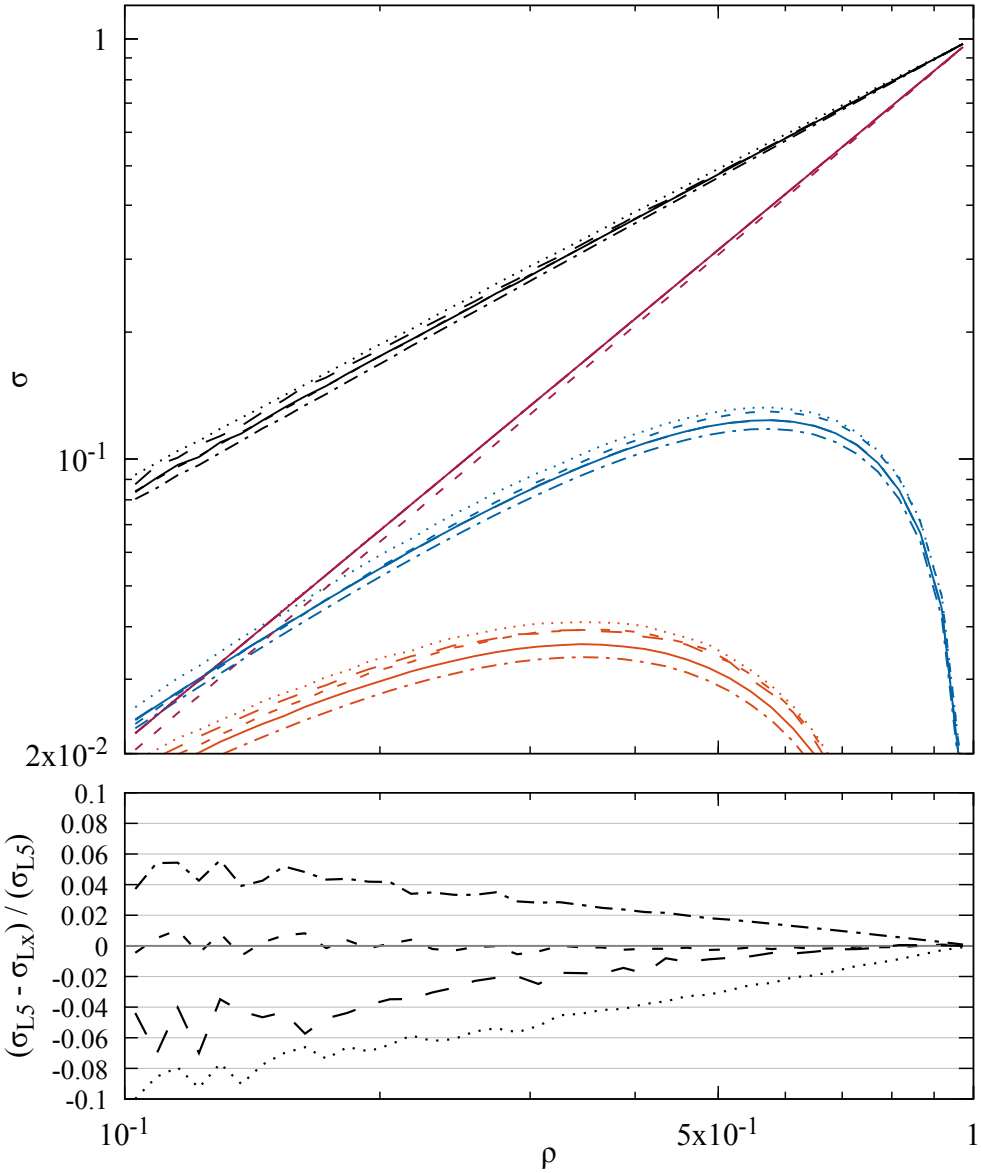


Figure 4.23: The veto cross-section for the t-channel gluon exchange contribution to $q\bar{q} \to q\bar{q}$ in the asymmetric configuration. Solid: Full colour (L5), Dash-dotted: LC′ + FCR (L4), Long-dashed: LC′ + LCR + singlets (L3), Dotted: LC′ + LCR (L2), Short-dashed: strict LC (L1). For the L1–L4 curves we start the evolution using the leading-colour approximation to the hard-scatter matrix.

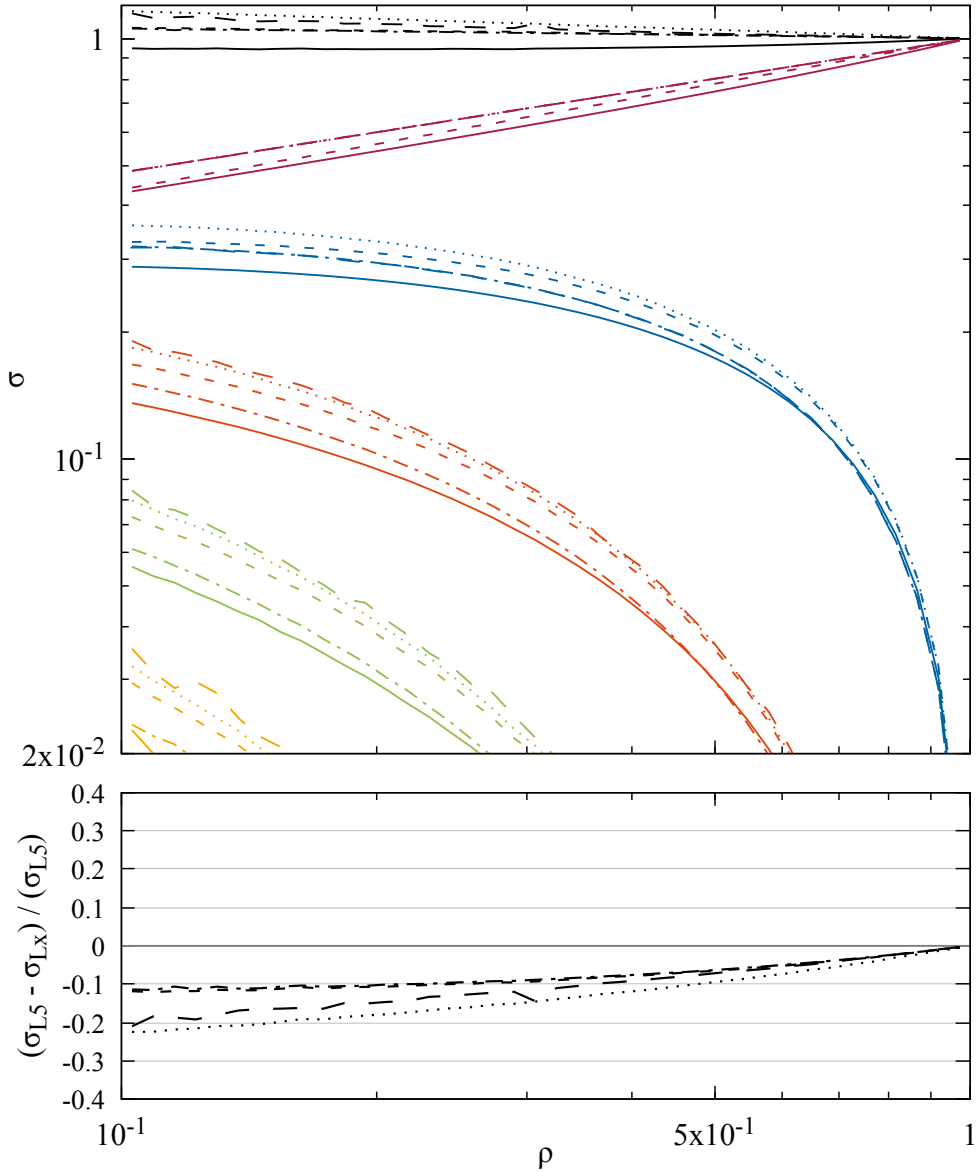


Figure 4.24: The veto cross-section for the s-channel gluon exchange contribution to $q\bar{q} \to q\bar{q}$ in the asymmetric configuration. Solid: Full colour (L5), Dash-dotted: LC′ + FCR (L4), Long-dashed: LC′ + LCR + singlets (L3), Dotted: LC′ + LCR (L2), Short-dashed: strict LC (L1). For the L1–L4 curves we start the evolution using the leading-colour approximation to the hard-scatter matrix.

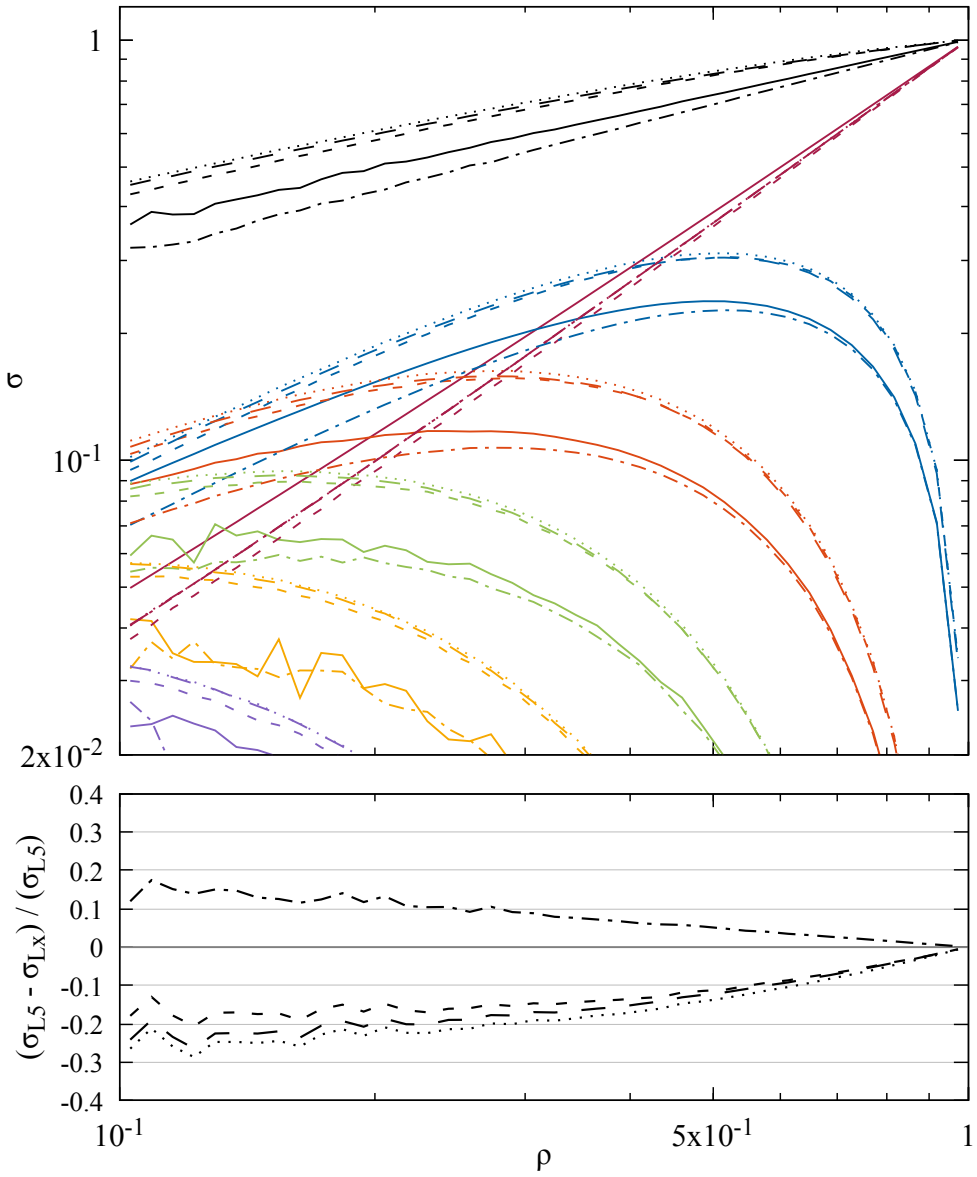


Figure 4.25: The veto cross-section for the s-channel quark exchange contribution to $qg \to qg$ in the symmetric configuration. Solid: Full colour (L5), Dash-dotted: LC′ + FCR (L4), Long-dashed: LC′ + LCR + singlets (L3), Dotted: LC′ + LCR (L2), Short-dashed: strict LC (L1). Evolution starts from the full-colour hard-scatter matrix.

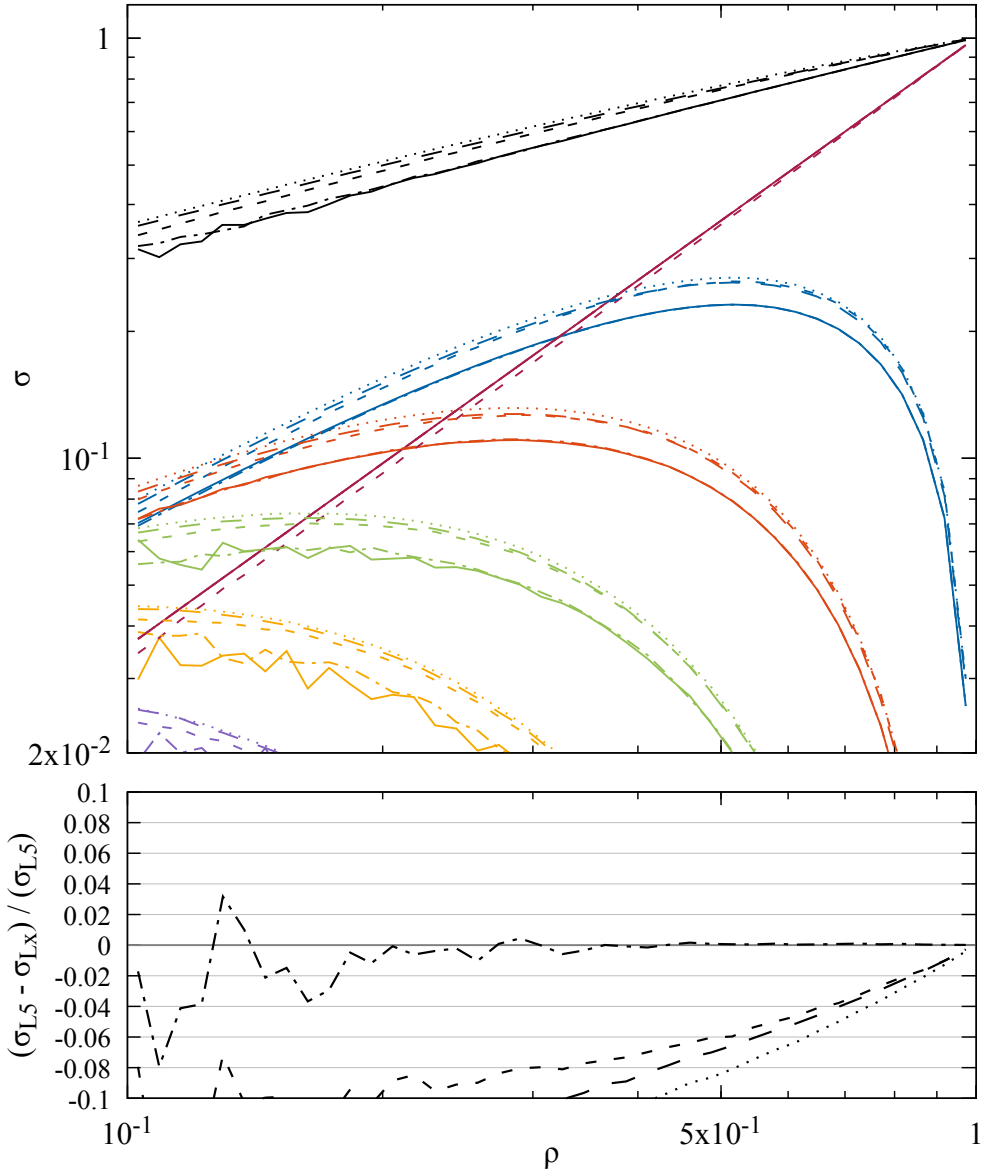


Figure 4.26: The veto cross-section for the t-channel gluon exchange contribution to $qg \to qg$ in the symmetric configuration. Solid: Full colour (L5), Dash-dotted: LC′ + FCR (L4), Long-dashed: LC′ + LCR + singlets (L3), Dotted: LC′ + LCR (L2), Short-dashed: strict LC (L1). Evolution starts from the full-colour hard-scatter matrix.

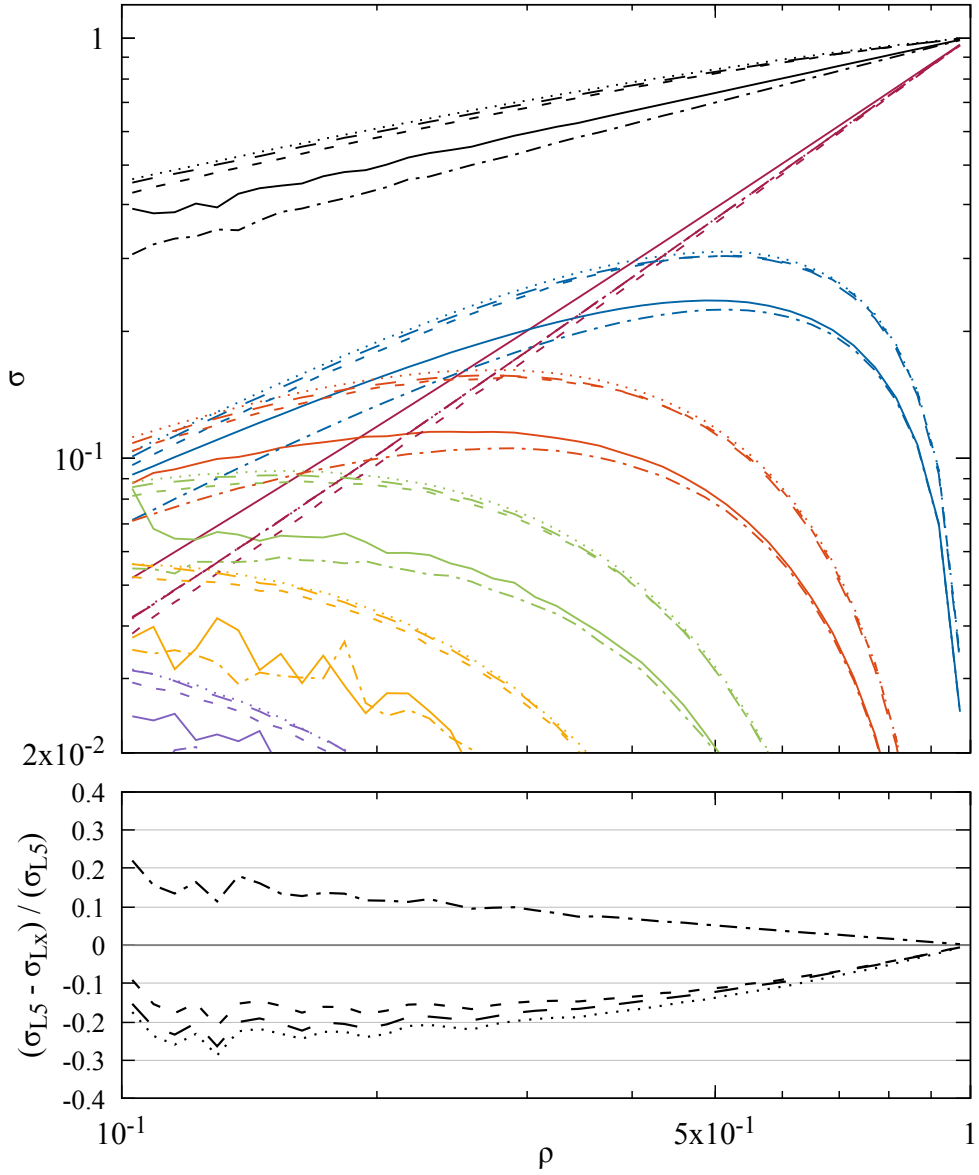


Figure 4.27: The veto cross-section for the u-channel quark exchange contribution to $qg \to qg$ in the symmetric configuration. Solid: Full colour (L5), Dash-dotted: LC′ + FCR (L4), Long-dashed: LC′ + LCR + singlets (L3), Dotted: LC′ + LCR (L2), Short-dashed: strict LC (L1). Evolution starts from the full-colour hard-scatter matrix.

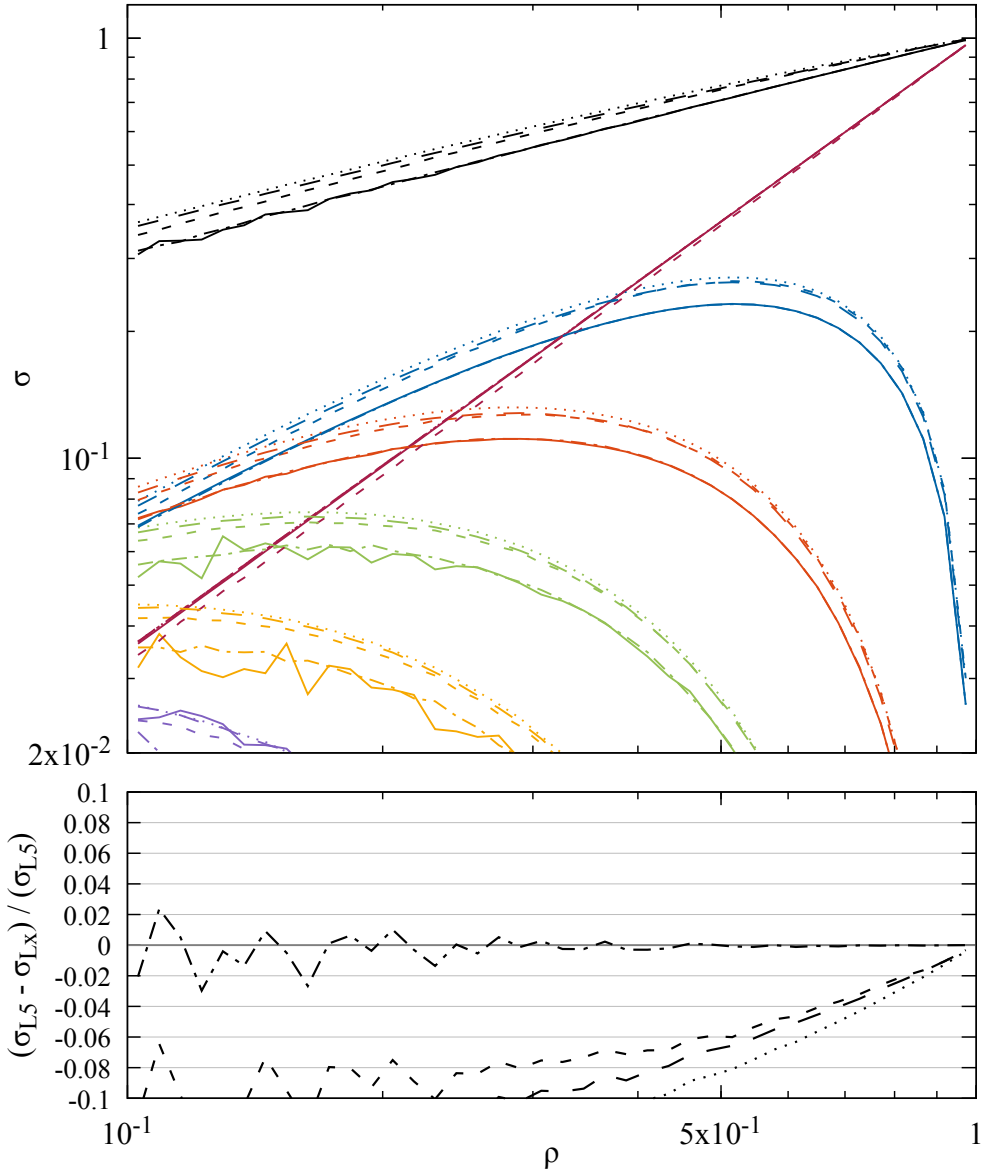


Figure 4.28: The veto cross-section for the st-channel interference contribution to $qg \to qg$ in the symmetric configuration. Solid: Full colour (L5), Dash-dotted: LC′ + FCR (L4), Long-dashed: LC′ + LCR + singlets (L3), Dotted: LC′ + LCR (L2), Short-dashed: strict LC (L1). Evolution starts from the full-colour hard-scatter matrix.

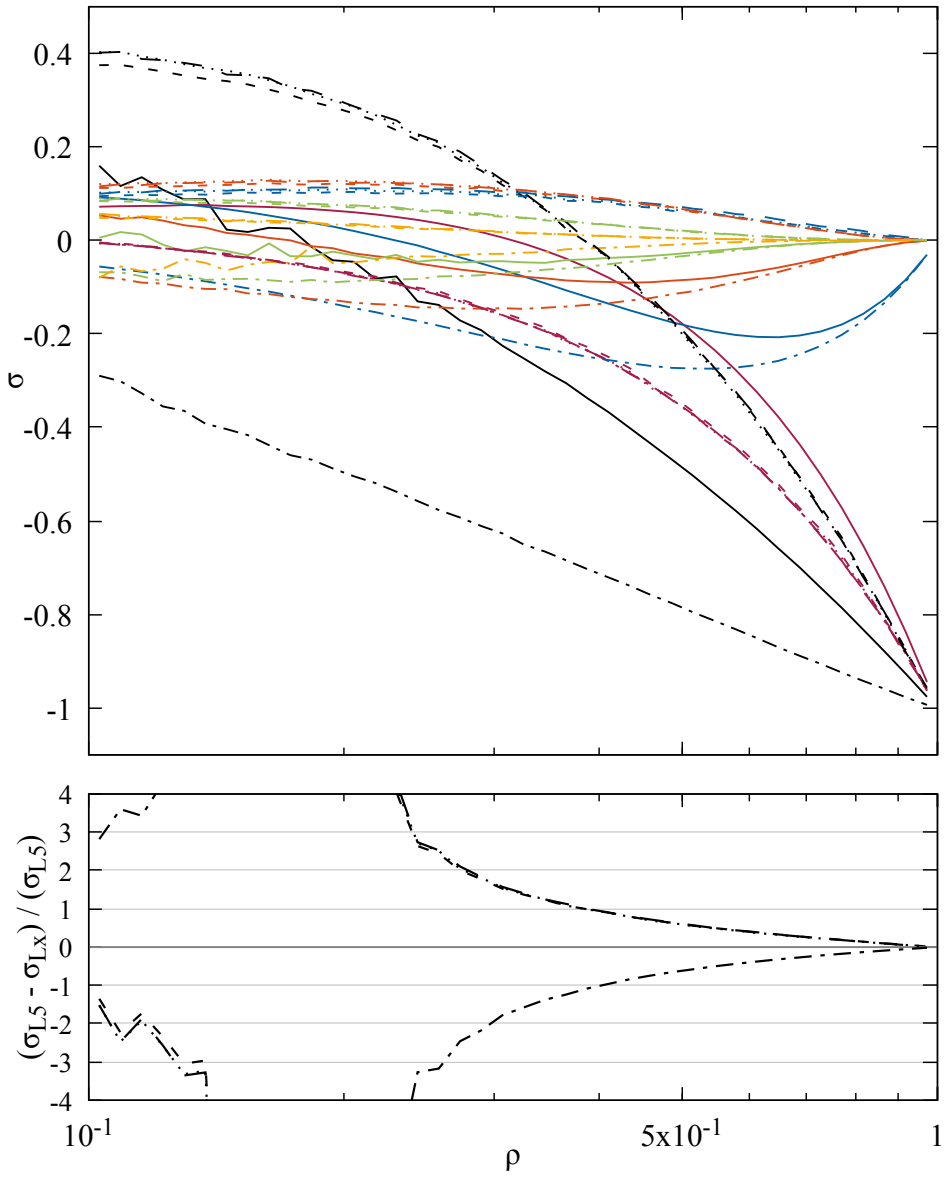


Figure 4.29: The veto cross-section for the su-channel interference contribution to $qg \to qg$ in the symmetric configuration. Solid: Full colour (L5), Dash-dotted: LC′ + FCR (L4), Long-dashed: LC′ + LCR + singlets (L3), Dotted: LC′ + LCR (L2), Short-dashed: strict LC (L1). Evolution starts from the full-colour hard-scatter matrix.

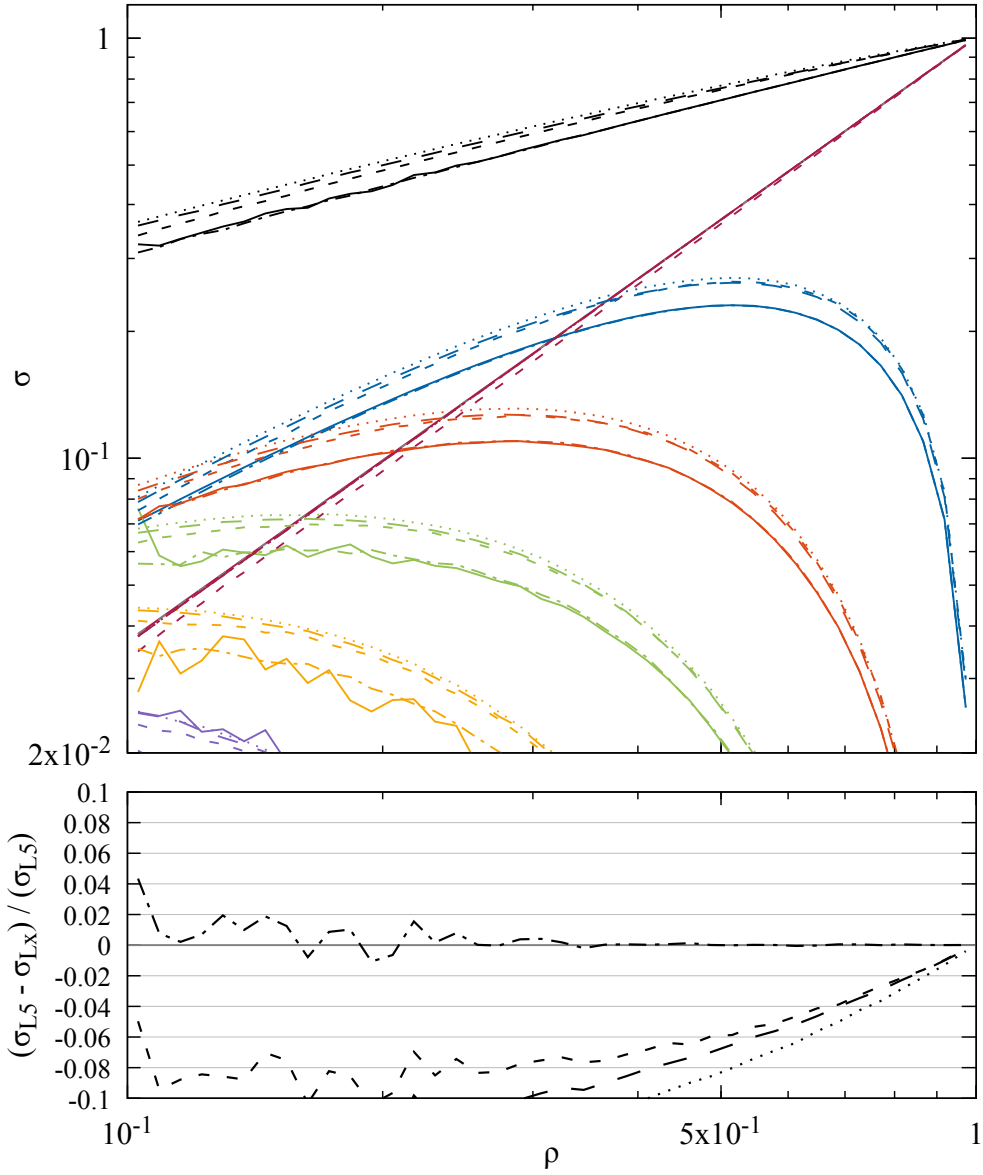


Figure 4.30: The veto cross-section for the tu-channel interference contribution to $qg \to qg$ in the symmetric configuration. Solid: Full colour (L5), Dash-dotted: LC′ + FCR (L4), Long-dashed: LC′ + LCR + singlets (L3), Dotted: LC′ + LCR (L2), Short-dashed: strict LC (L1). Evolution starts from the full-colour hard-scatter matrix.

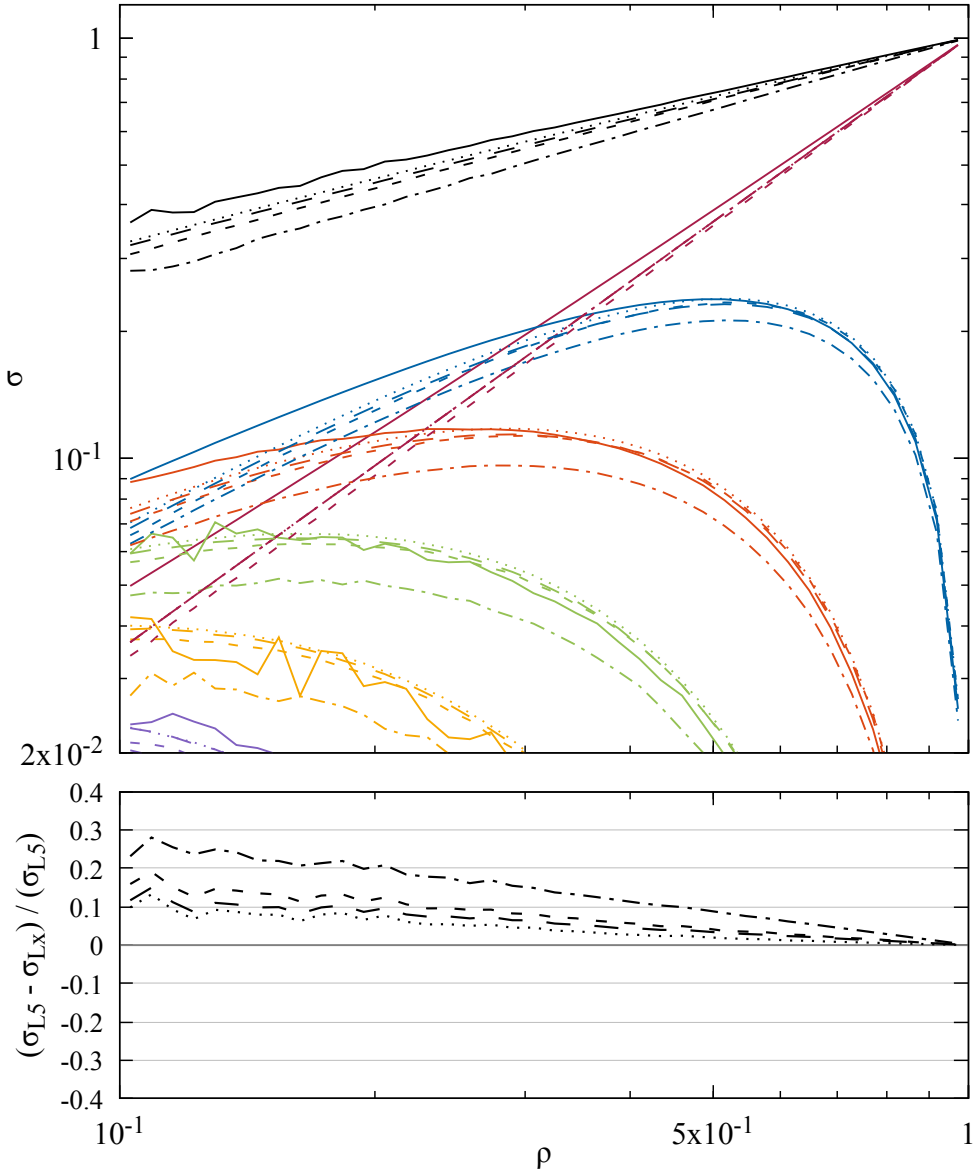


Figure 4.31: The veto cross-section for the s-channel quark exchange contribution to $qg \to qg$ in the symmetric configuration. Solid: Full colour (L5), Dash-dotted: LC′ + FCR (L4), Long-dashed: LC′ + LCR + singlets (L3), Dotted: LC′ + LCR (L2), Short-dashed: strict LC (L1). Evolution starts from the leading-colour hard-scatter matrix.

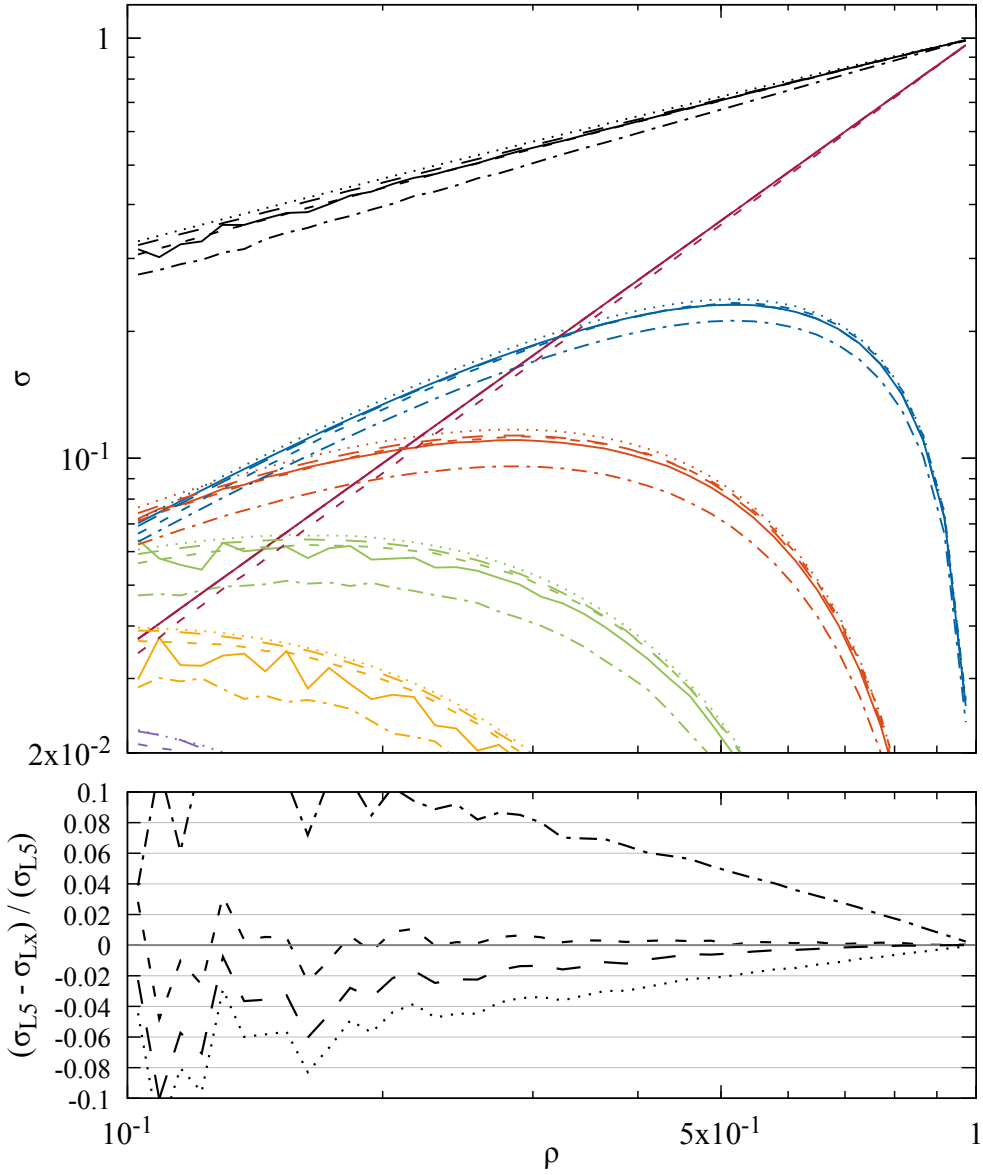


Figure 4.32: The veto cross-section for the t-channel gluon exchange contribution to $qg \to qg$ in the symmetric configuration. Solid: Full colour (L5), Dash-dotted: LC′ + FCR (L4), Long-dashed: LC′ + LCR + singlets (L3), Dotted: LC′ + LCR (L2), Short-dashed: strict LC (L1). Evolution starts from the leading-colour hard-scatter matrix.

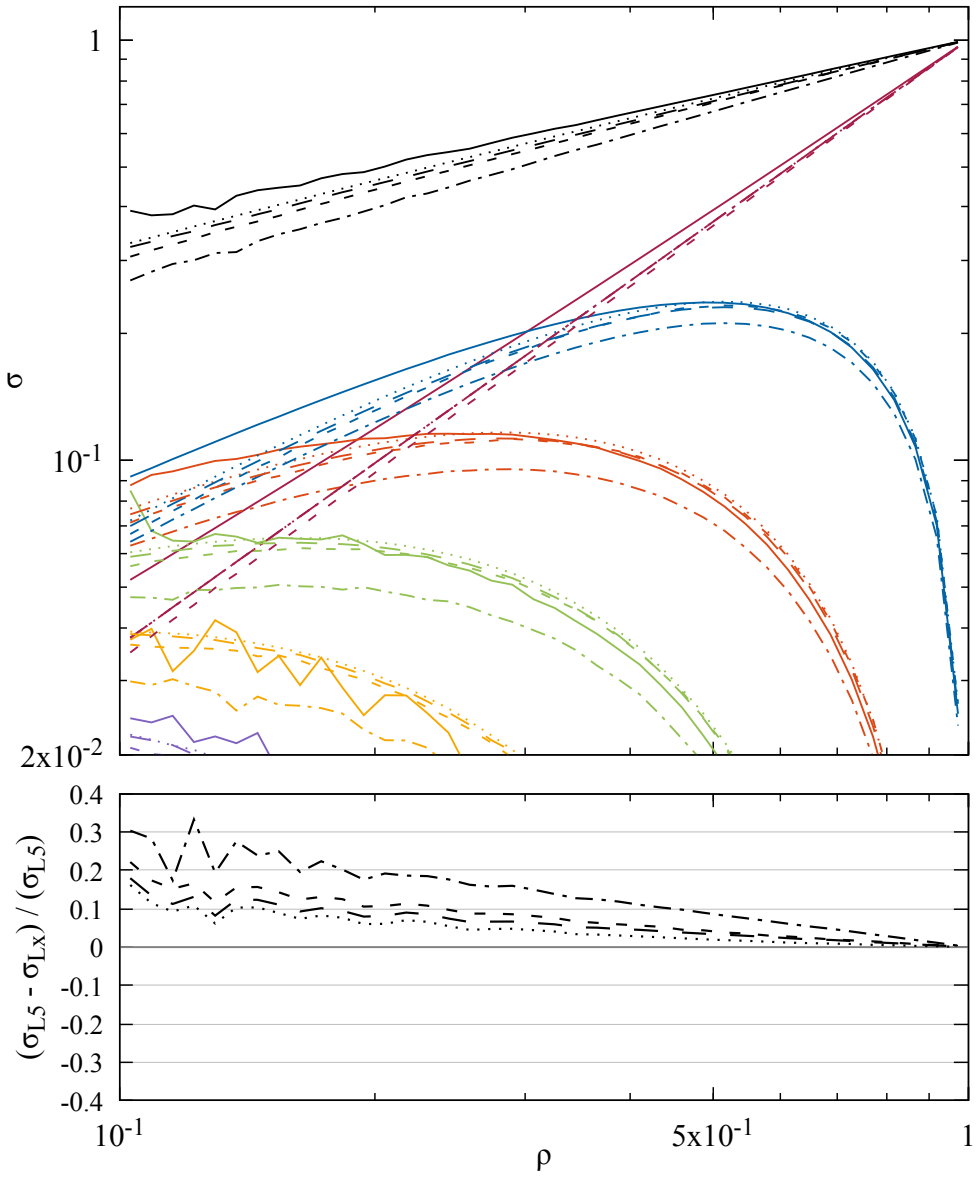


Figure 4.33: The veto cross-section for the u-channel quark exchange contribution to $qg \to qg$ in the symmetric configuration. Solid: Full colour (L5), Dash-dotted: LC′ + FCR (L4), Long-dashed: LC′ + LCR + singlets (L3), Dotted: LC′ + LCR (L2), Short-dashed: strict LC (L1). Evolution starts from the leading-colour hard-scatter matrix.

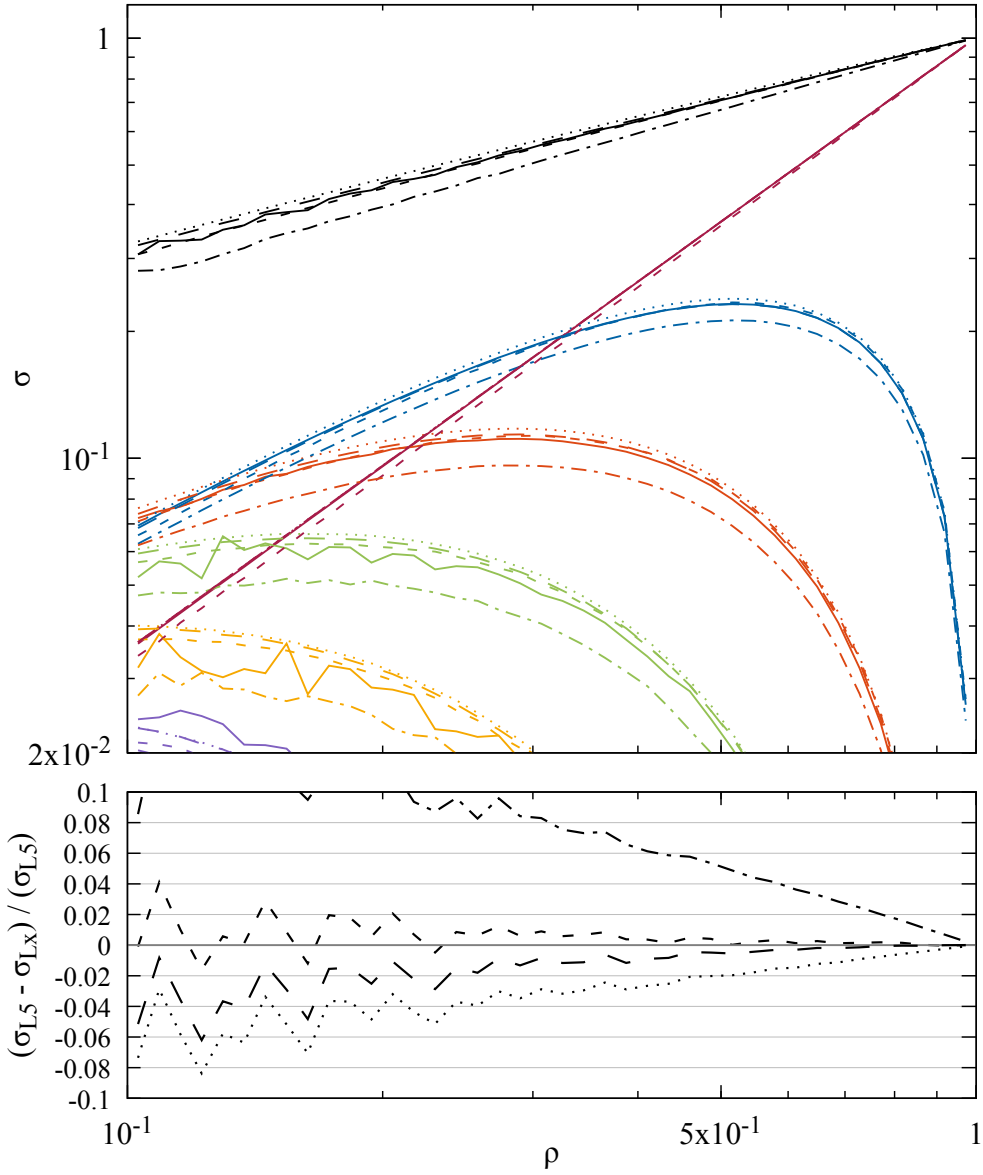


Figure 4.34: The veto cross-section for the st-channel interference contribution to $qg \to qg$ in the symmetric configuration. Solid: Full colour (L5), Dash-dotted: LC′ + FCR (L4), Long-dashed: LC′ + LCR + singlets (L3), Dotted: LC′ + LCR (L2), Short-dashed: strict LC (L1). Evolution starts from the leading-colour hard-scatter matrix.

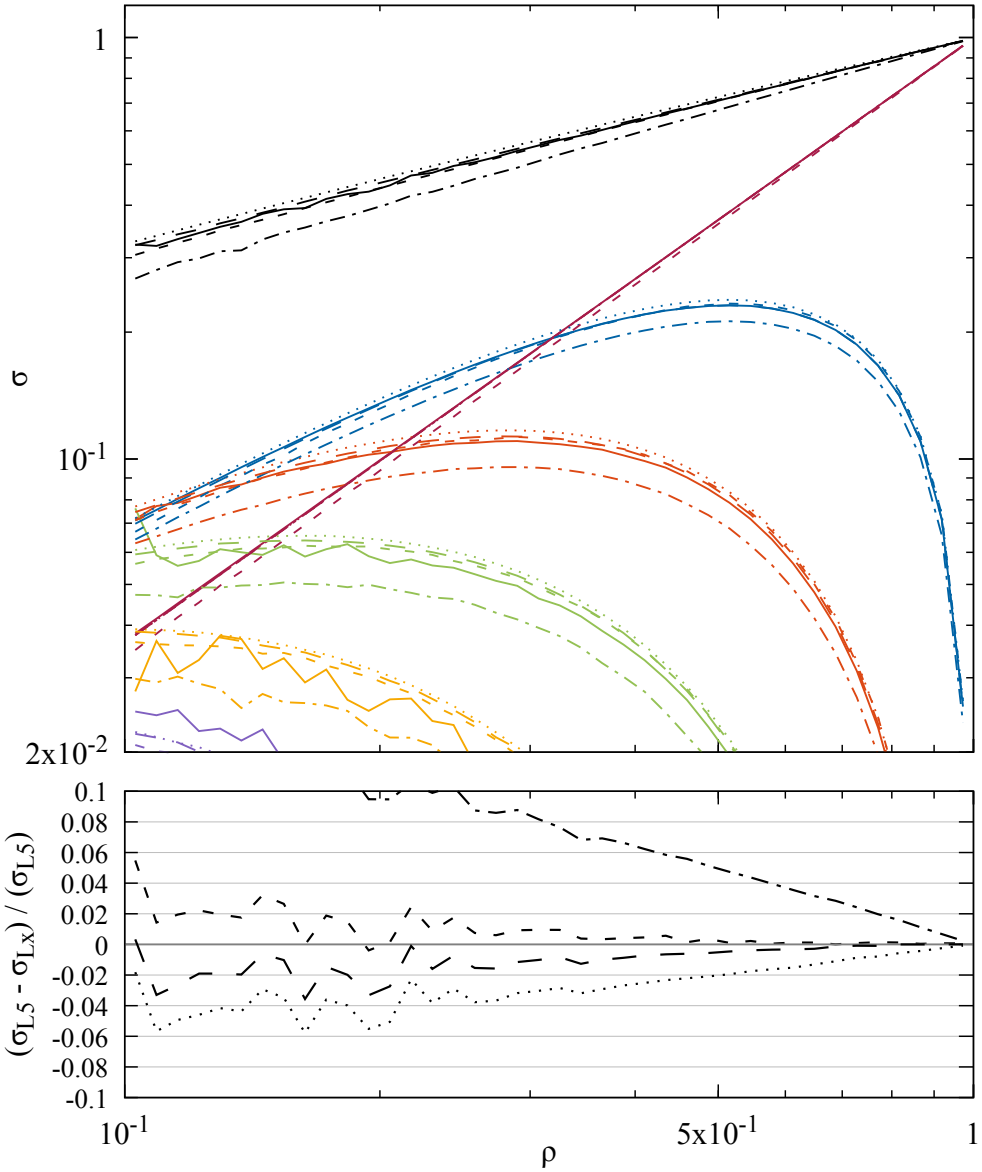


Figure 4.35: The veto cross-section for the tu-channel interference contribution to $qg \to qg$ in the symmetric configuration. Solid: Full colour (L5), Dash-dotted: LC′ + FCR (L4), Long-dashed: LC′ + LCR + singlets (L3), Dotted: LC′ + LCR (L2), Short-dashed: strict LC (L1). Evolution starts from the leading-colour hard-scatter matrix.

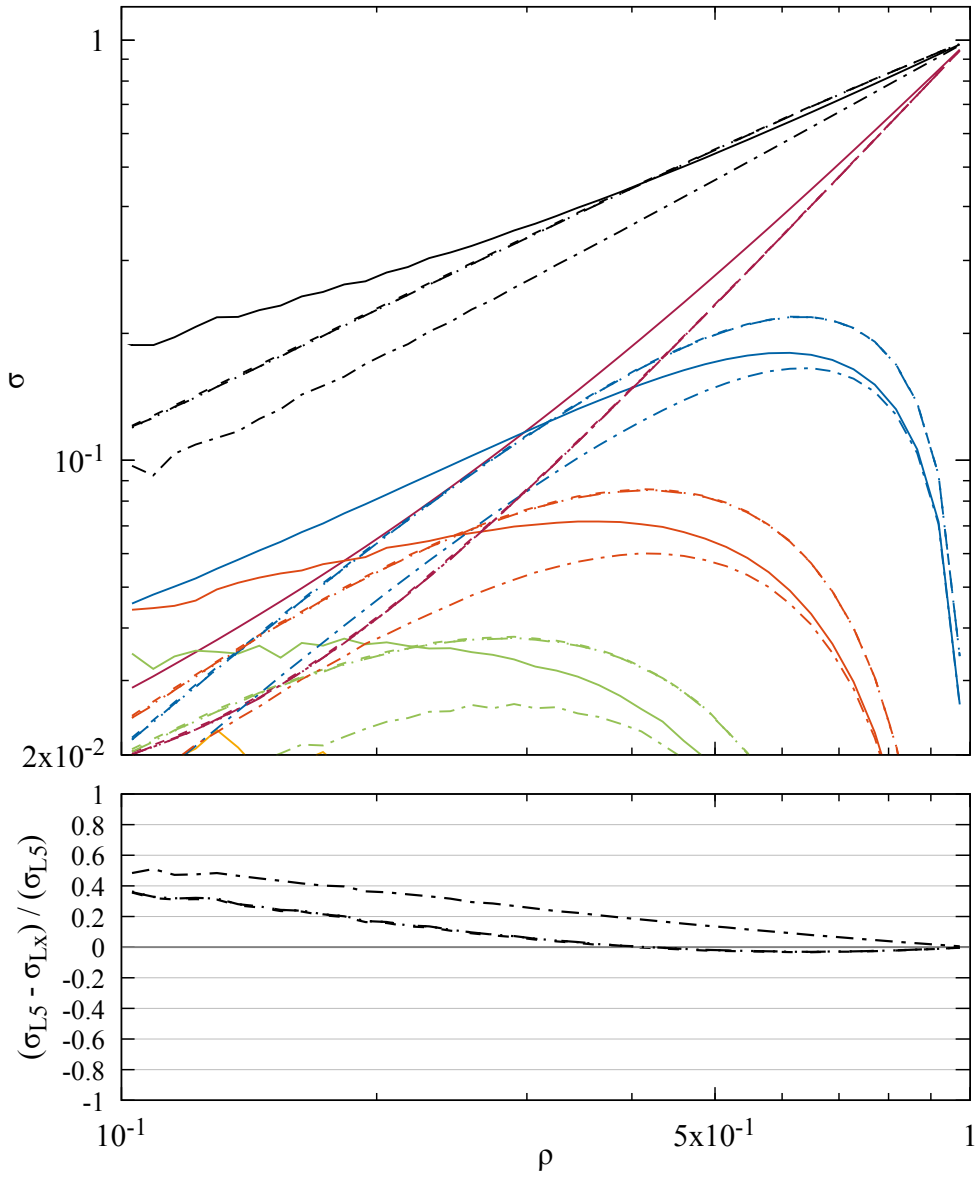


Figure 4.36: The veto cross-section for the s-channel quark exchange contribution to $qg \to qg$ in the asymmetric configuration. Solid: Full colour (L5), Dash-dotted: LC′ + FCR (L4), Long-dashed: LC′ + LCR + singlets (L3), Dotted: LC′ + LCR (L2), Short-dashed: strict LC (L1). Evolution starts from the full-colour hard-scatter matrix.

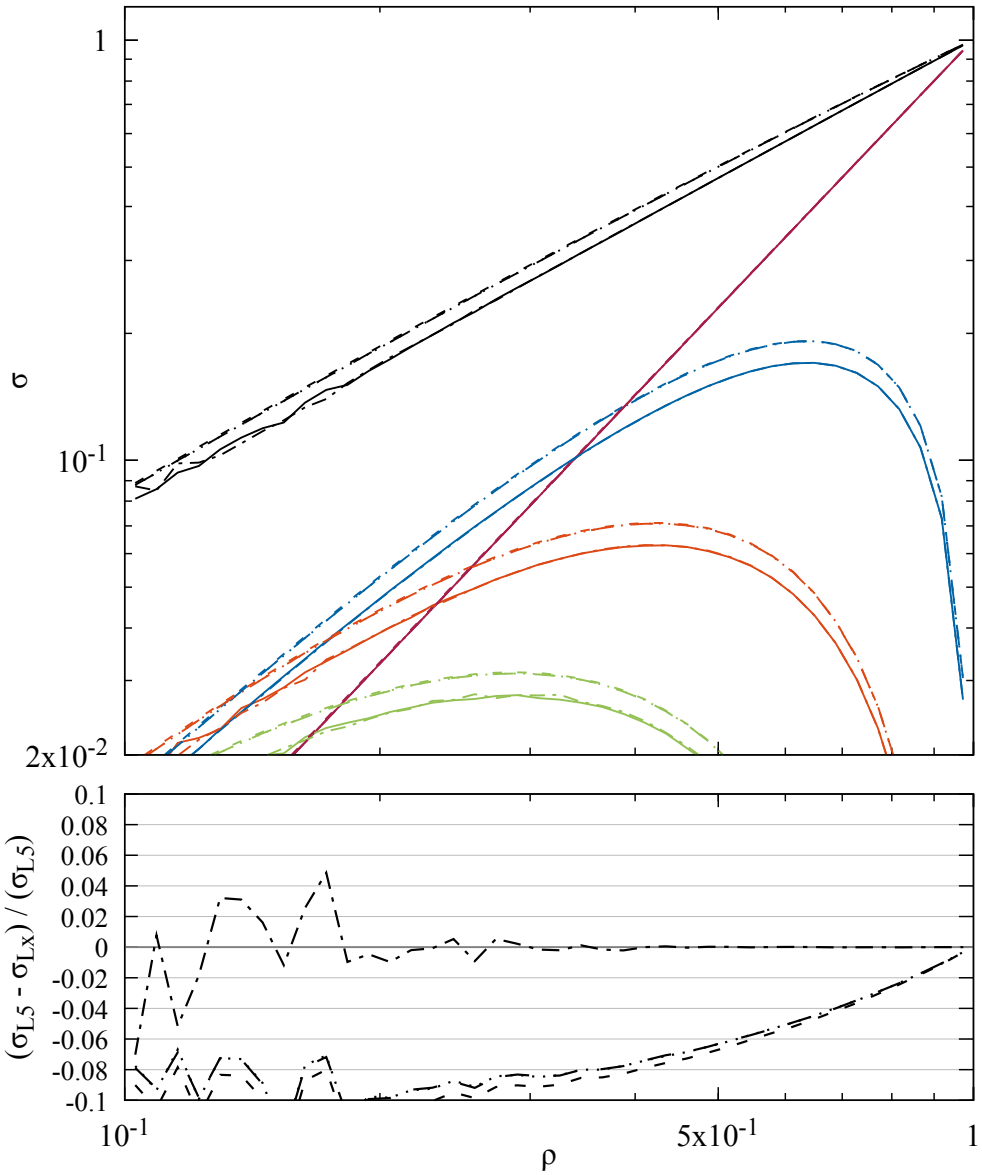


Figure 4.37: The veto cross-section for the t-channel gluon exchange contribution to $qg \to qg$ in the asymmetric configuration. Solid: Full colour (L5), Dash-dotted: LC′ + FCR (L4), Long-dashed: LC′ + LCR + singlets (L3), Dotted: LC′ + LCR (L2), Short-dashed: strict LC (L1). Evolution starts from the full-colour hard-scatter matrix.

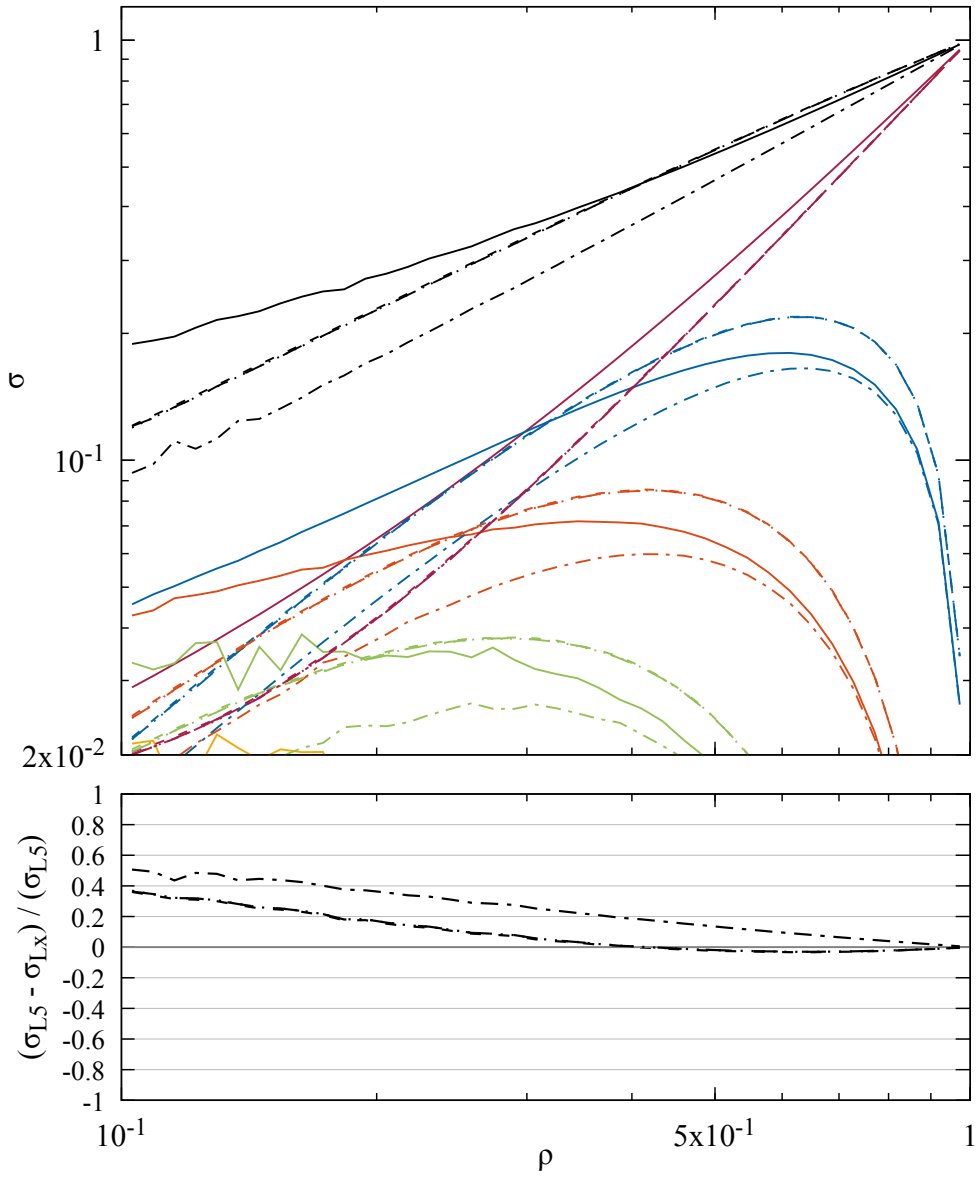


Figure 4.38: The veto cross-section for the u-channel quark exchange contribution to $qg \to qg$ in the asymmetric configuration. Solid: Full colour (L5), Dash-dotted: LC′ + FCR (L4), Long-dashed: LC′ + LCR + singlets (L3), Dotted: LC′ + LCR (L2), Short-dashed: strict LC (L1). Evolution starts from the full-colour hard-scatter matrix.

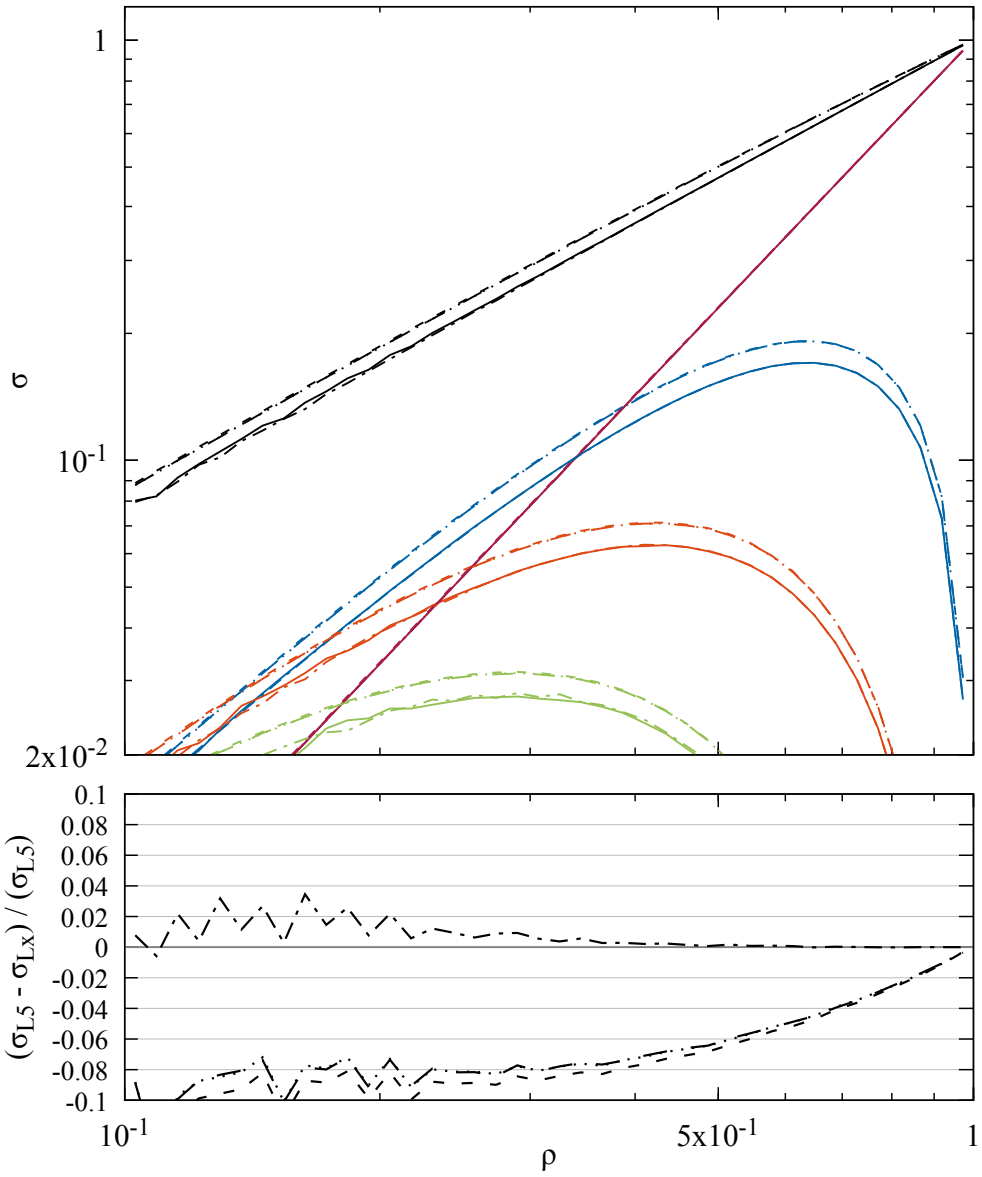


Figure 4.39: The veto cross-section for the st-channel interference contribution to $qg \to qg$ in the asymmetric configuration. Solid: Full colour (L5), Dash-dotted: LC′ + FCR (L4), Long-dashed: LC′ + LCR + singlets (L3), Dotted: LC′ + LCR (L2), Short-dashed: strict LC (L1). Evolution starts from the full-colour hard-scatter matrix.

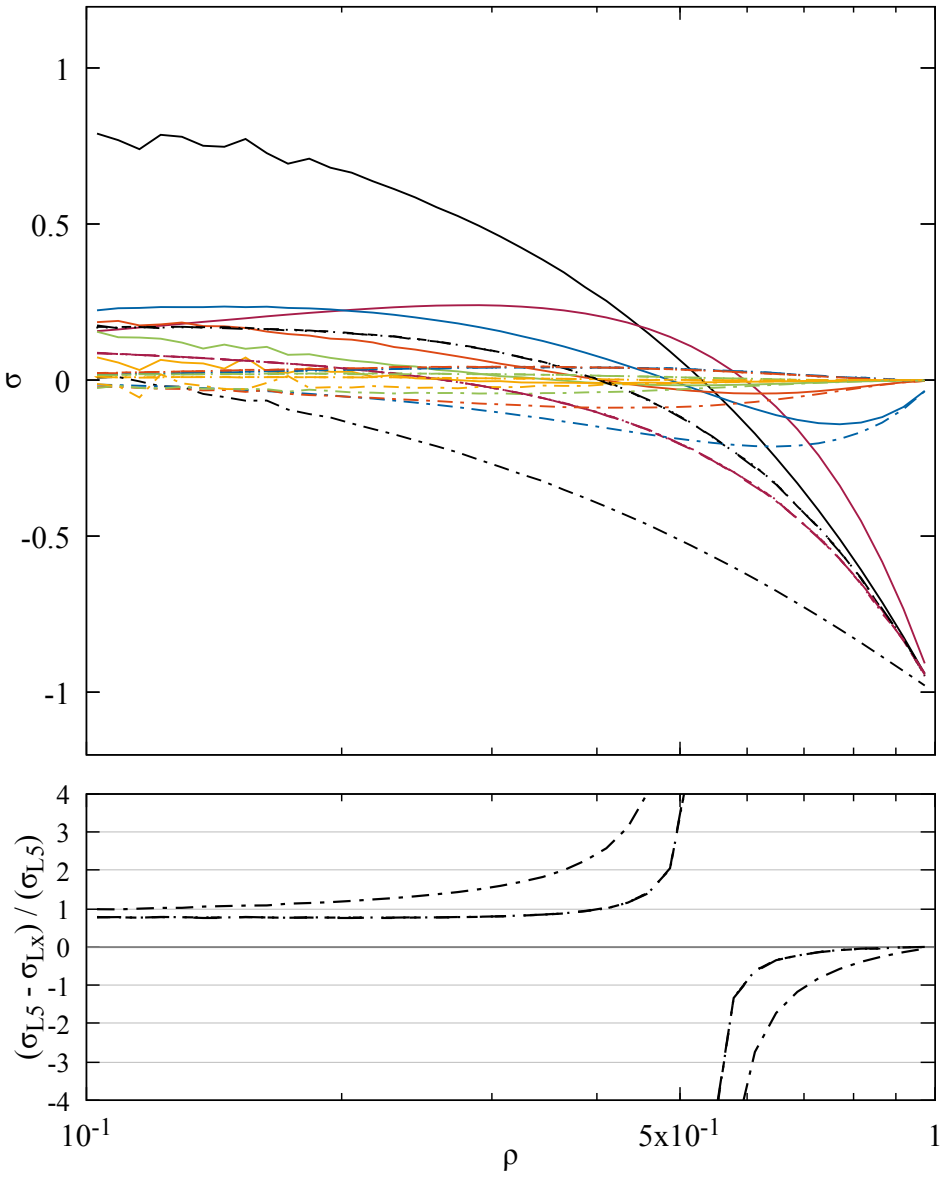


Figure 4.40: The veto cross-section for the su-channel interference contribution to $qg \to qg$ in the asymmetric configuration. Solid: Full colour (L5), Dash-dotted: LC′ + FCR (L4), Long-dashed: LC′ + LCR + singlets (L3), Dotted: LC′ + LCR (L2), Short-dashed: strict LC (L1). Evolution starts from the full-colour hard-scatter matrix.

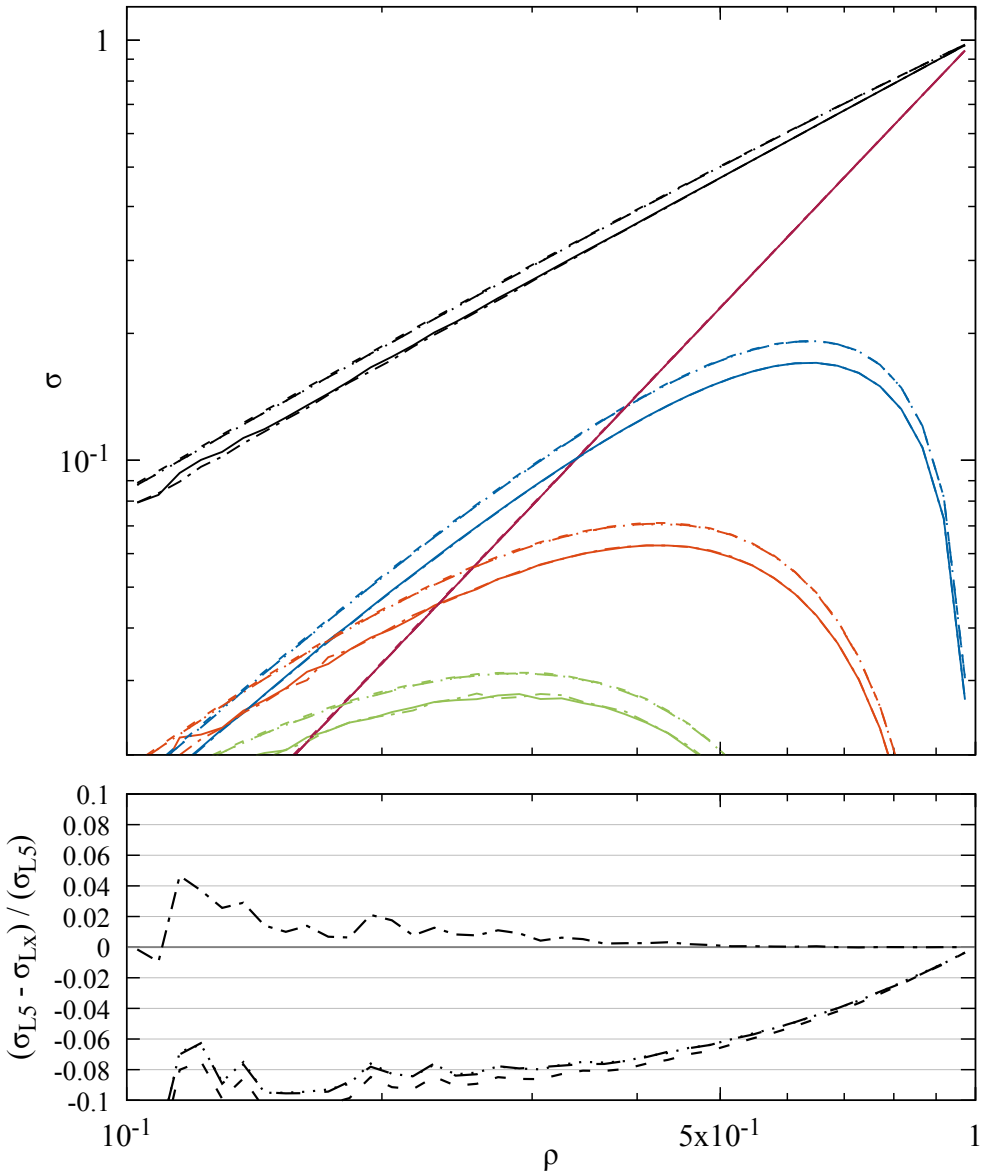


Figure 4.41: The veto cross-section for the tu-channel interference contribution to $qg \to qg$ in the asymmetric configuration. Solid: Full colour (L5), Dash-dotted: LC′ + FCR (L4), Long-dashed: LC′ + LCR + singlets (L3), Dotted: LC′ + LCR (L2), Short-dashed: strict LC (L1). Evolution starts from the full-colour hard-scatter matrix.

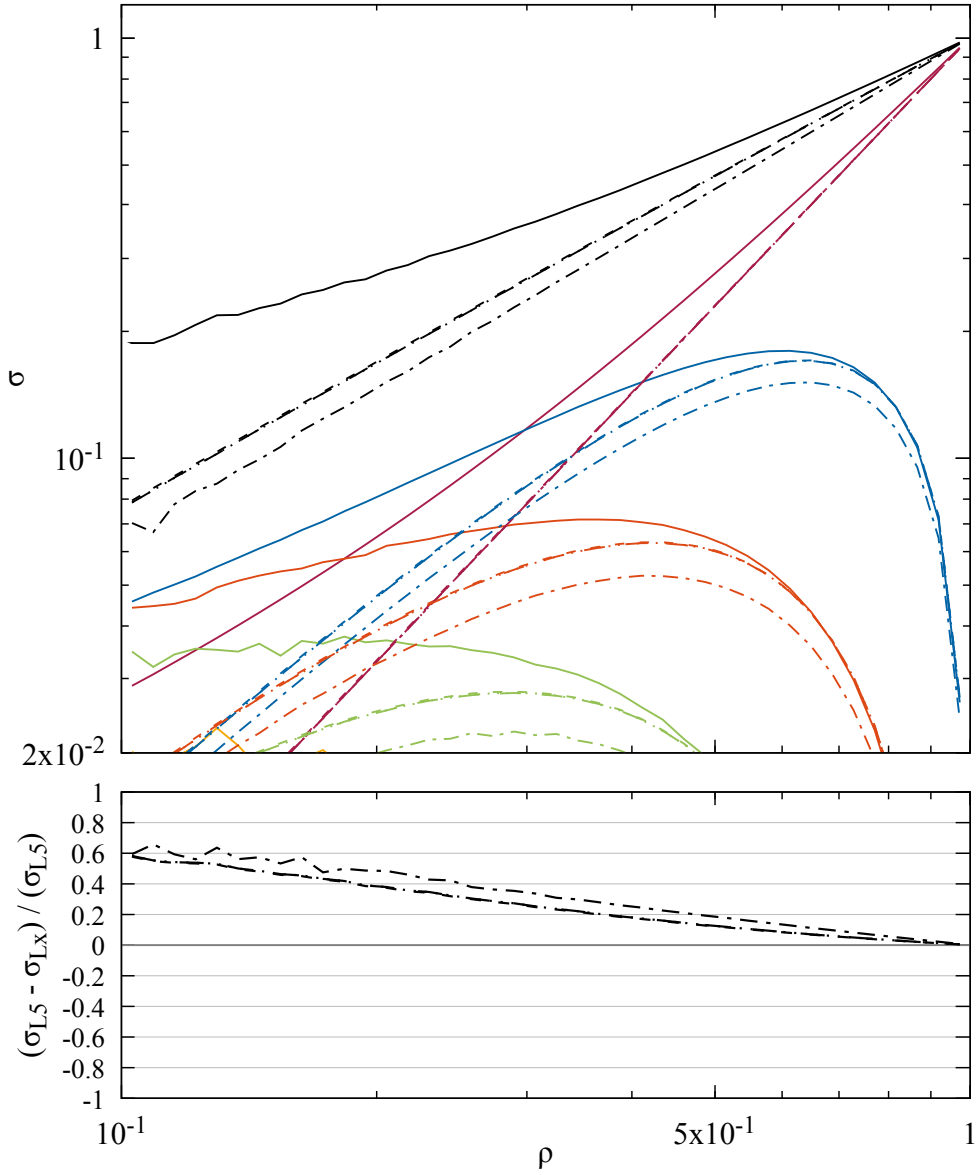


Figure 4.42: The veto cross-section for the s-channel quark exchange contribution to $qg \to qg$ in the asymmetric configuration. Solid: Full colour (L5), Dash-dotted: LC′ + FCR (L4), Long-dashed: LC′ + LCR + singlets (L3), Dotted: LC′ + LCR (L2), Short-dashed: strict LC (L1). Evolution starts from the leading-colour hard-scatter matrix.

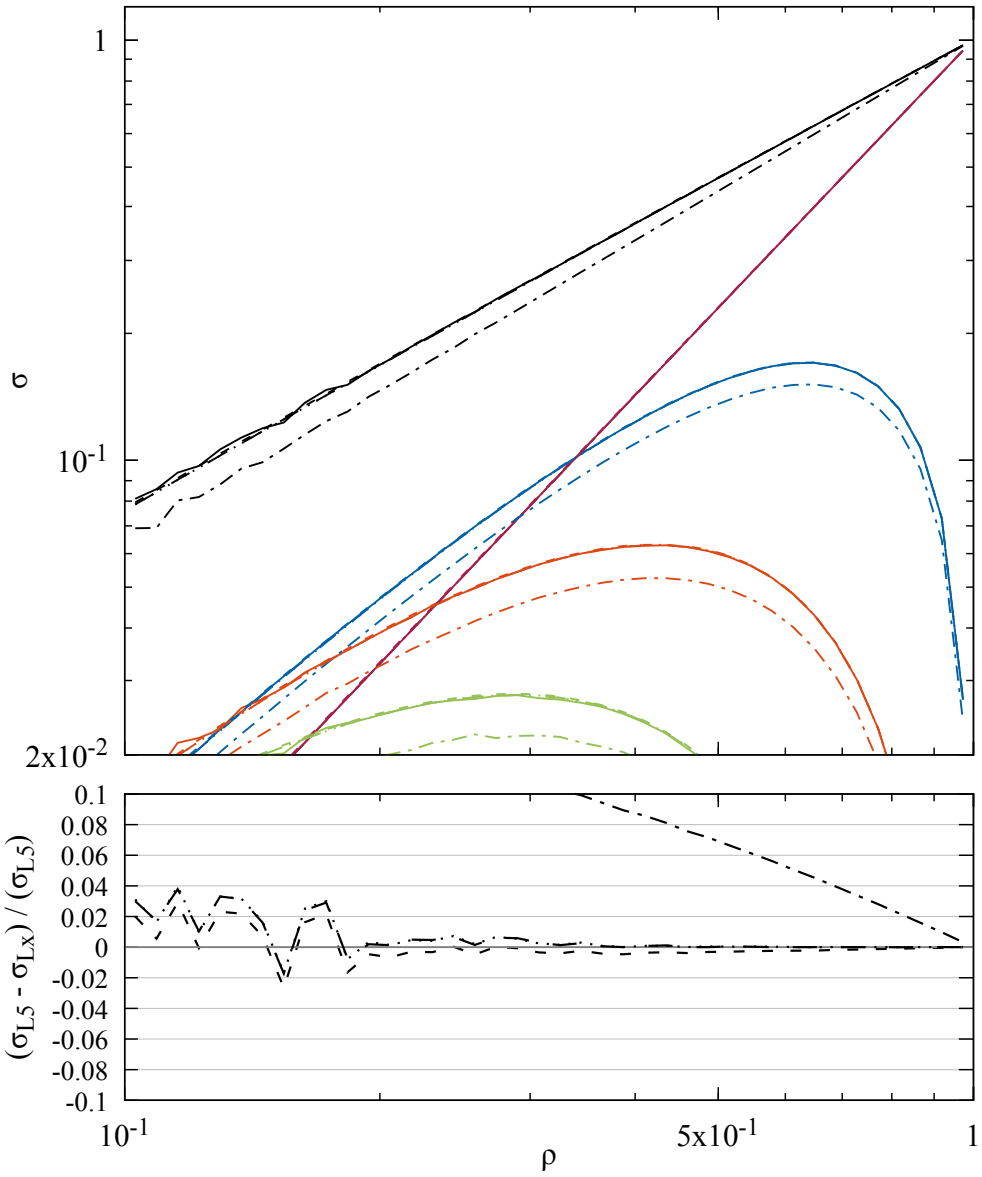


Figure 4.43: The veto cross-section for the t-channel gluon exchange contribution to $qg \to qg$ in the asymmetric configuration. Solid: Full colour (L5), Dash-dotted: LC′ + FCR (L4), Long-dashed: LC′ + LCR + singlets (L3), Dotted: LC′ + LCR (L2), Short-dashed: strict LC (L1). Evolution starts from the leading-colour hard-scatter matrix.

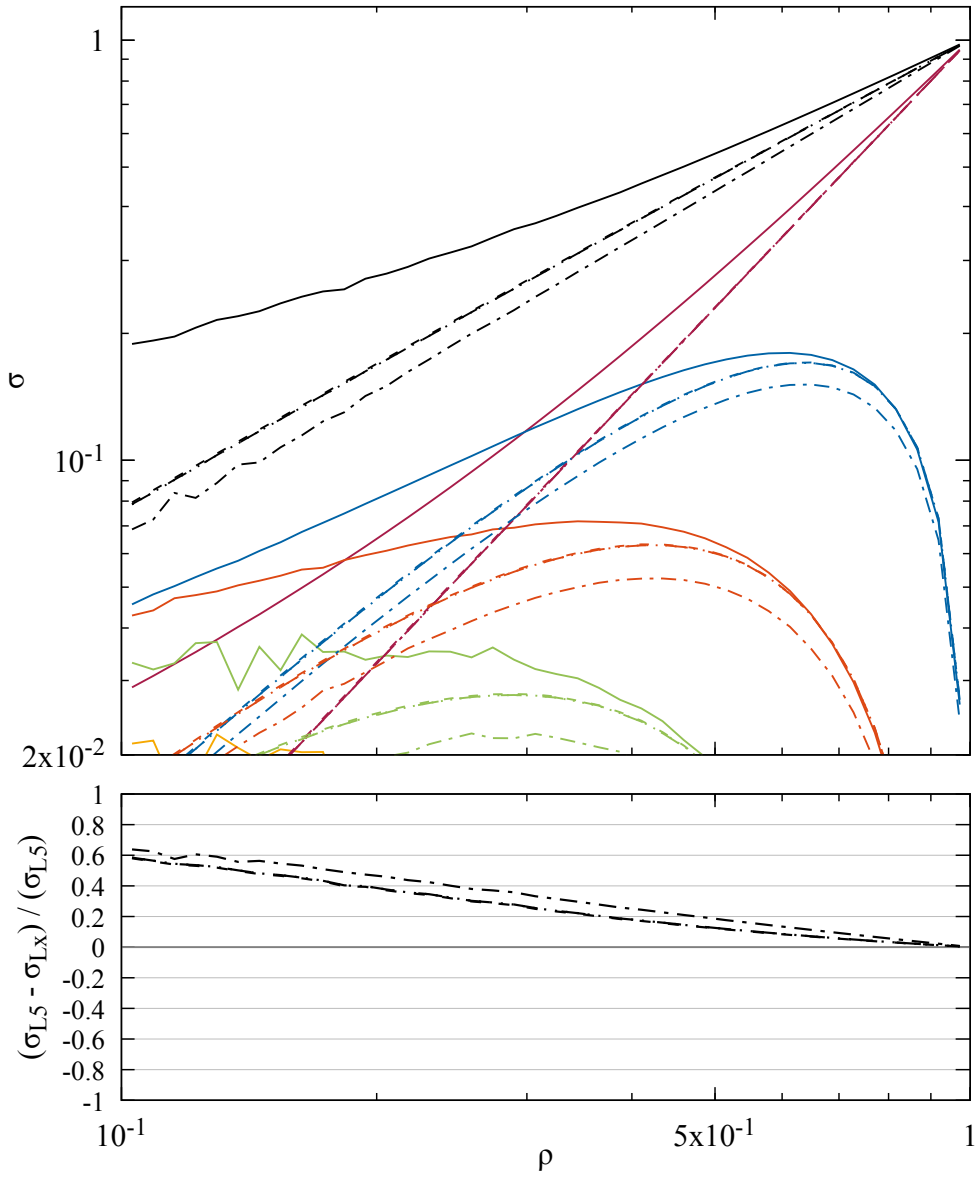


Figure 4.44: The veto cross-section for the u-channel quark exchange contribution to $qg \to qg$ in the asymmetric configuration. Solid: Full colour (L5), Dash-dotted: LC′ + FCR (L4), Long-dashed: LC′ + LCR + singlets (L3), Dotted: LC′ + LCR (L2), Short-dashed: strict LC (L1). Evolution starts from the leading-colour hard-scatter matrix.

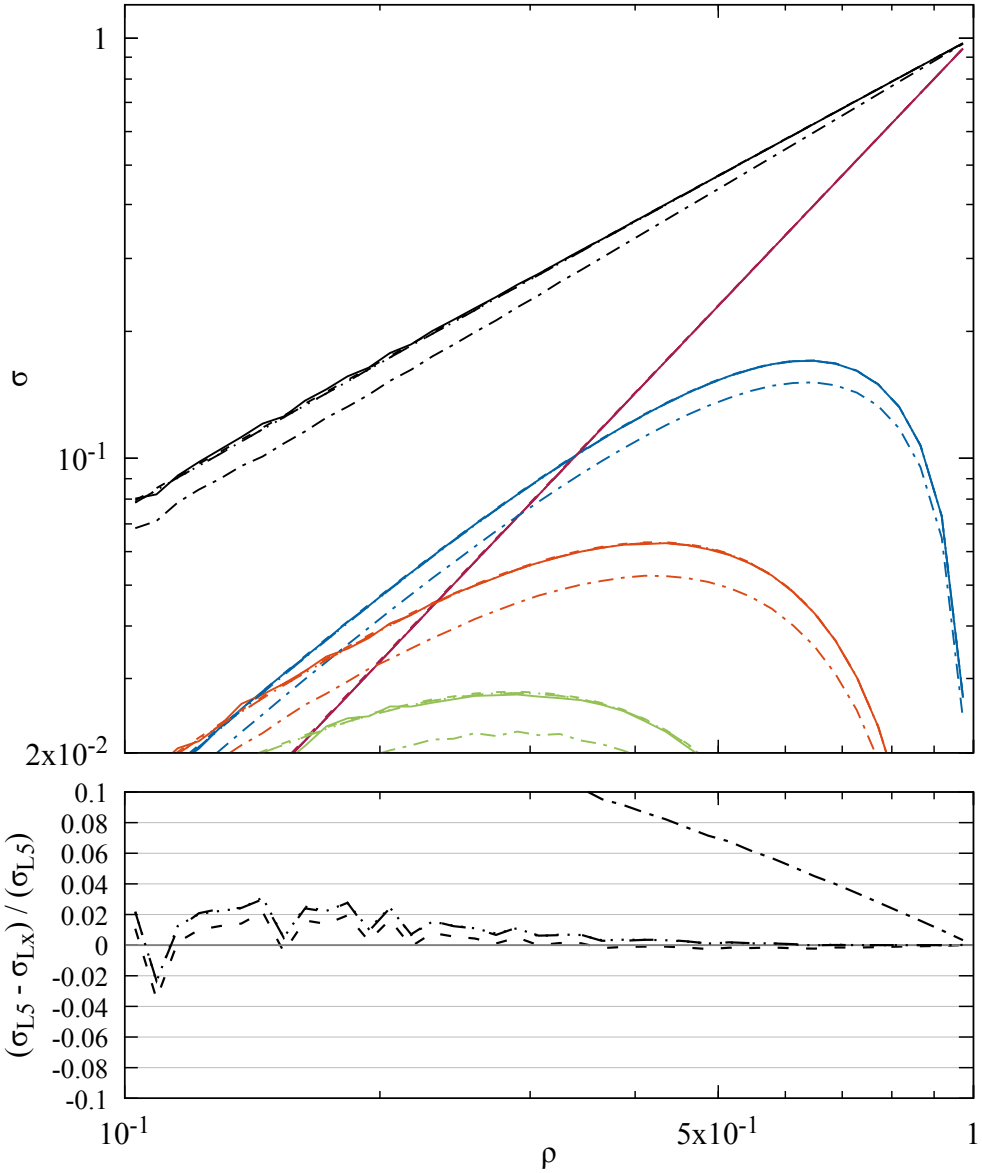


Figure 4.45: The veto cross-section for the st-channel interference contribution to $qg \to qg$ in the asymmetric configuration. Solid: Full colour (L5), Dash-dotted: LC′ + FCR (L4), Long-dashed: LC′ + LCR + singlets (L3), Dotted: LC′ + LCR (L2), Short-dashed: strict LC (L1). Evolution starts from the leading-colour hard-scatter matrix.

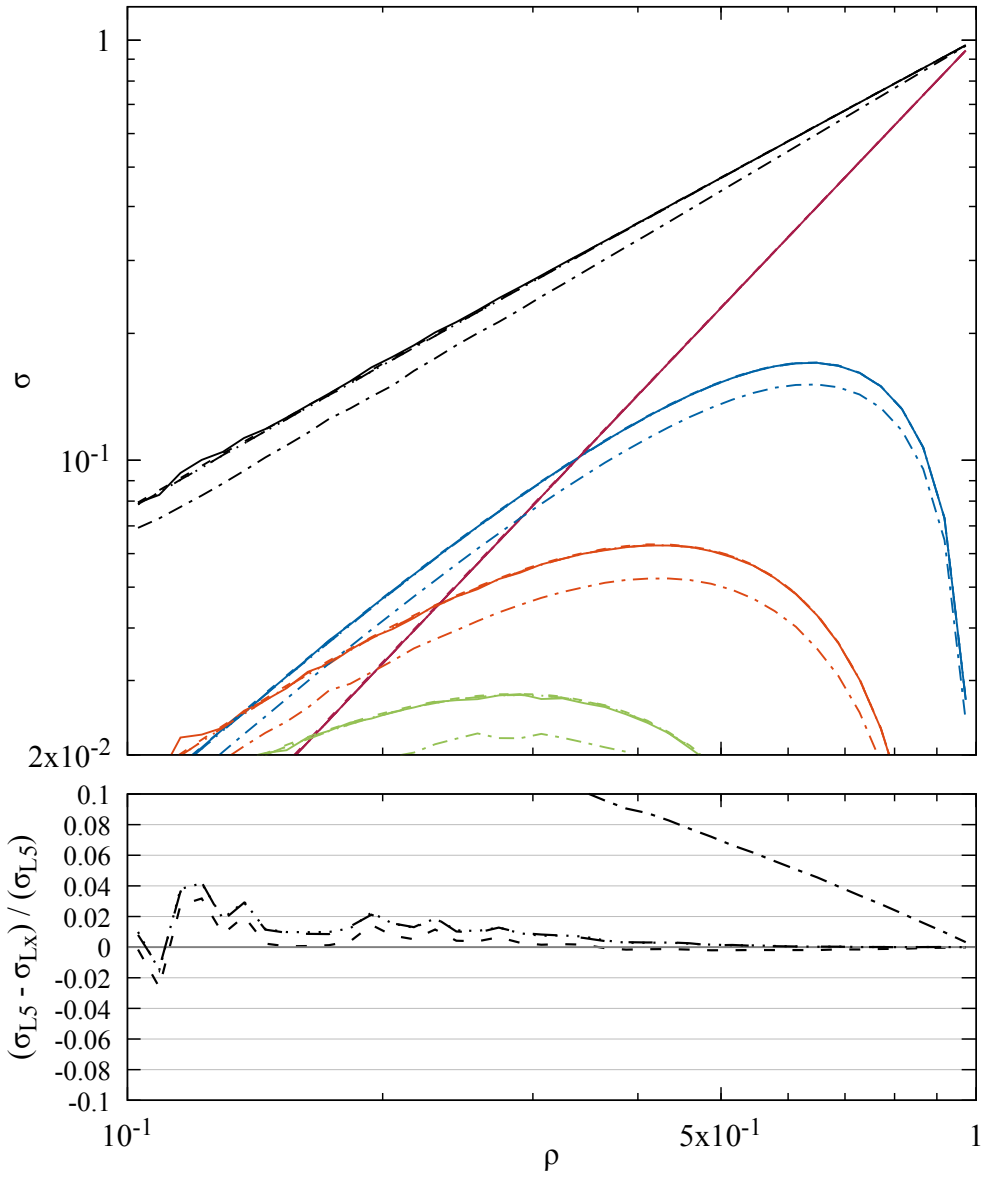


Figure 4.46: The veto cross-section for the tu-channel interference contribution to $qg \to qg$ in the asymmetric configuration. Solid: Full colour (L5), Dash-dotted: LC′ + FCR (L4), Long-dashed: LC′ + LCR + singlets (L3), Dotted: LC′ + LCR (L2), Short-dashed: strict LC (L1). Evolution starts from the leading-colour hard-scatter matrix.

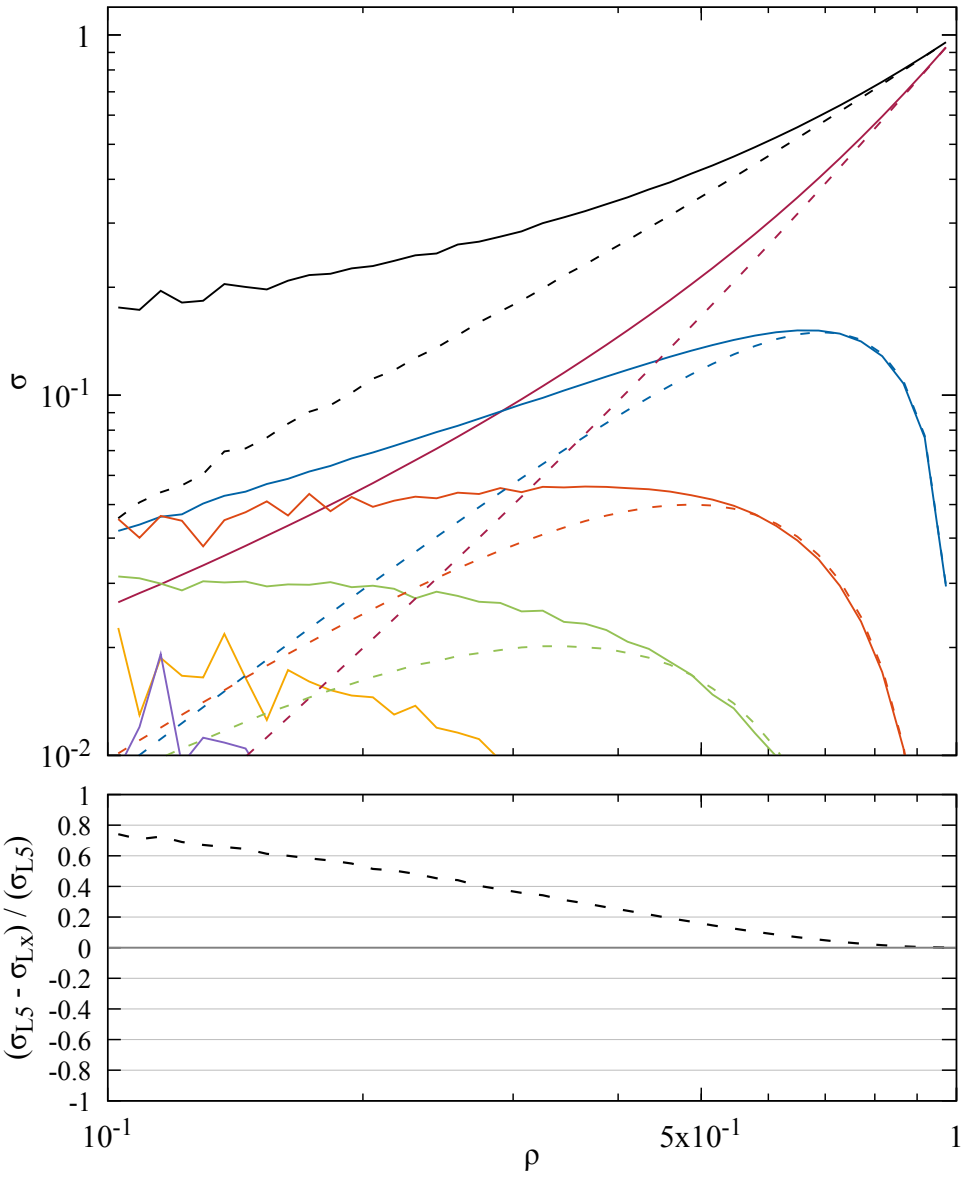


Figure 4.47: The veto cross-section for the s-channel contribution to $gg \to gg$. Solid: Full colour, Short-dashed: Leading colour.

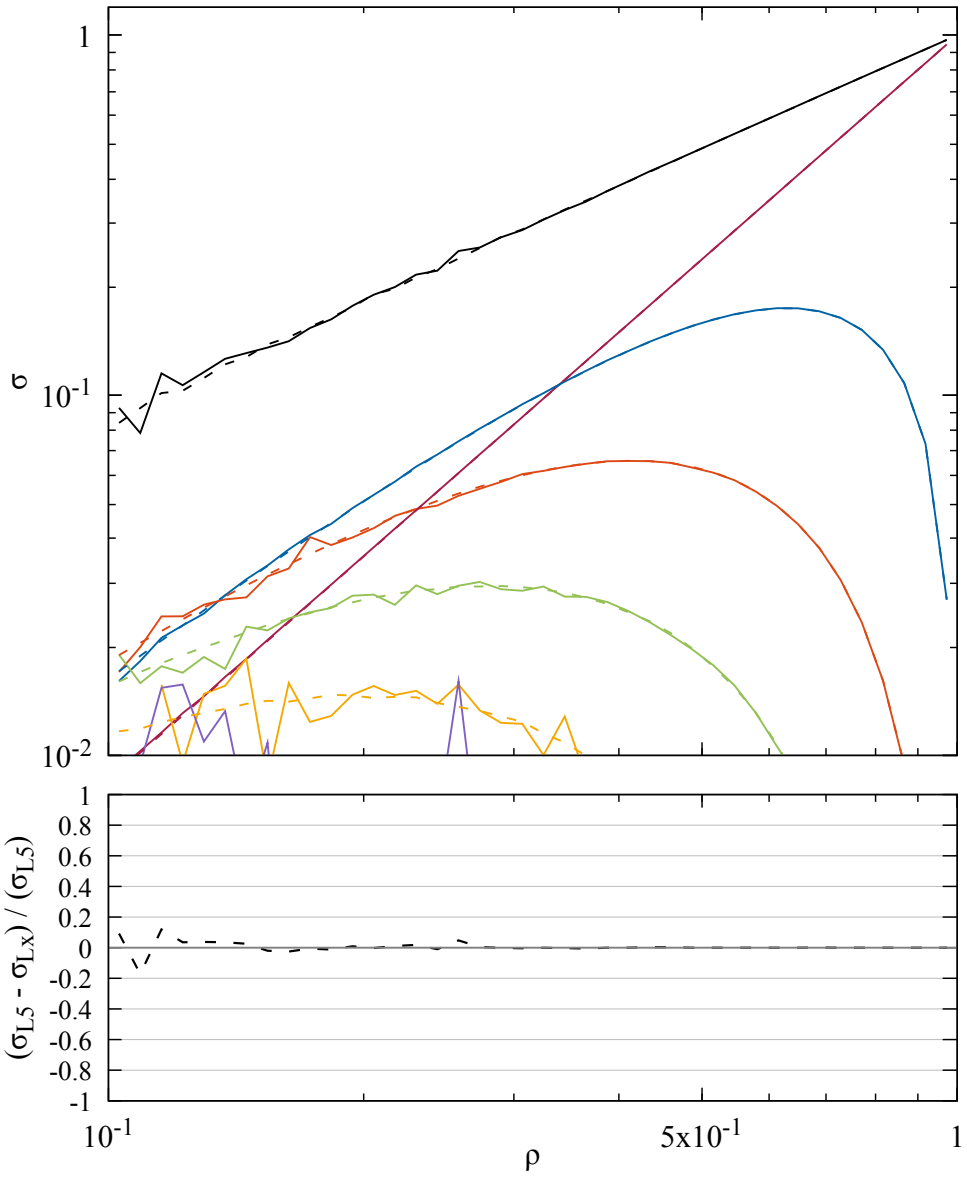


Figure 4.48: The veto cross-section for the t-channel contribution to $gg \to gg$. Solid: Full colour, Short-dashed: Leading colour.

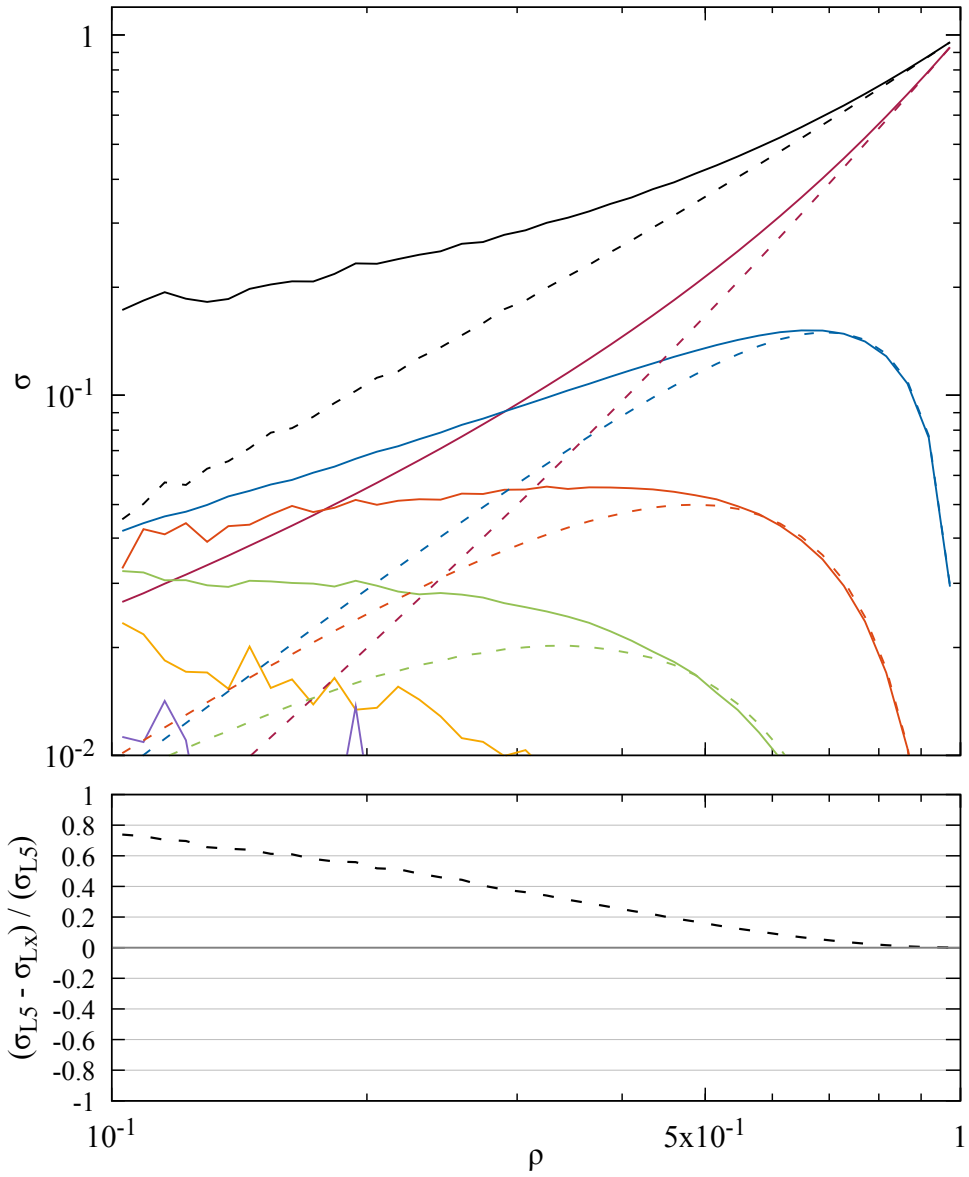


Figure 4.49: The veto cross-section for the u-channel contribution to $gg \to gg$. Solid: Full colour, Short-dashed: Leading colour.

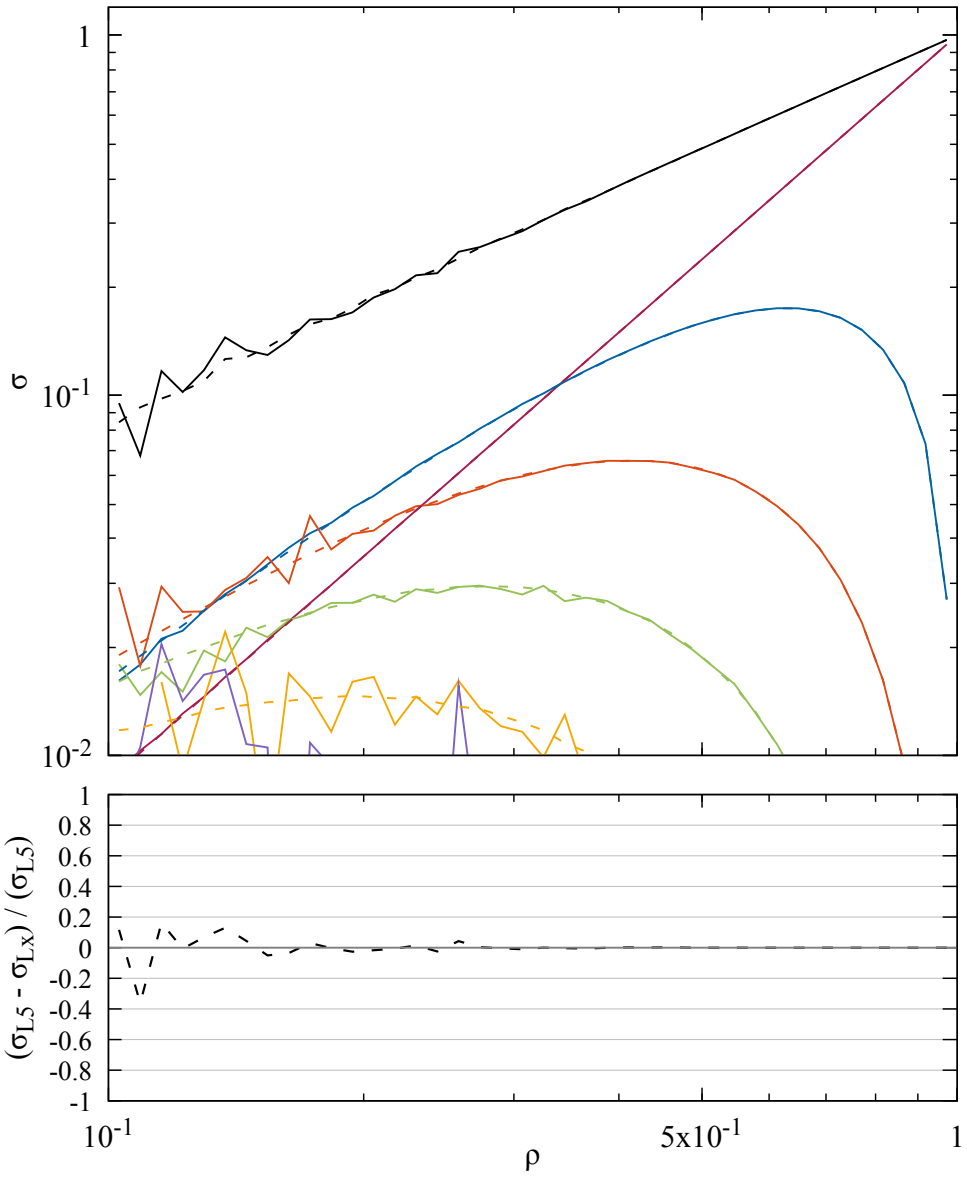


Figure 4.50: The veto cross-section for the st-interference contribution to $gg \to gg$. Solid: Full colour, Short-dashed: Leading colour.

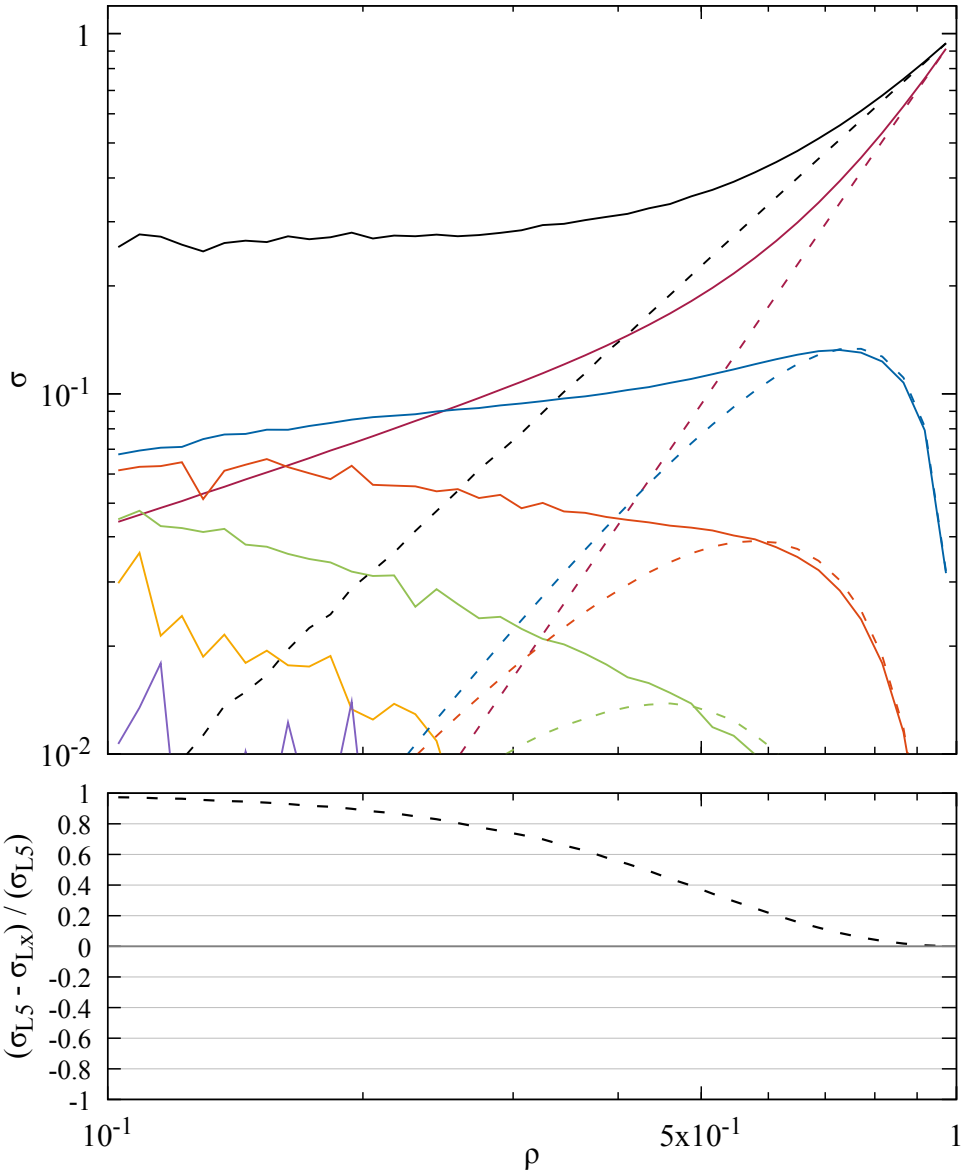


Figure 4.51: The veto cross-section for the su-interference contribution to $gg \to gg$. Solid: Full colour, Short-dashed: Leading colour.

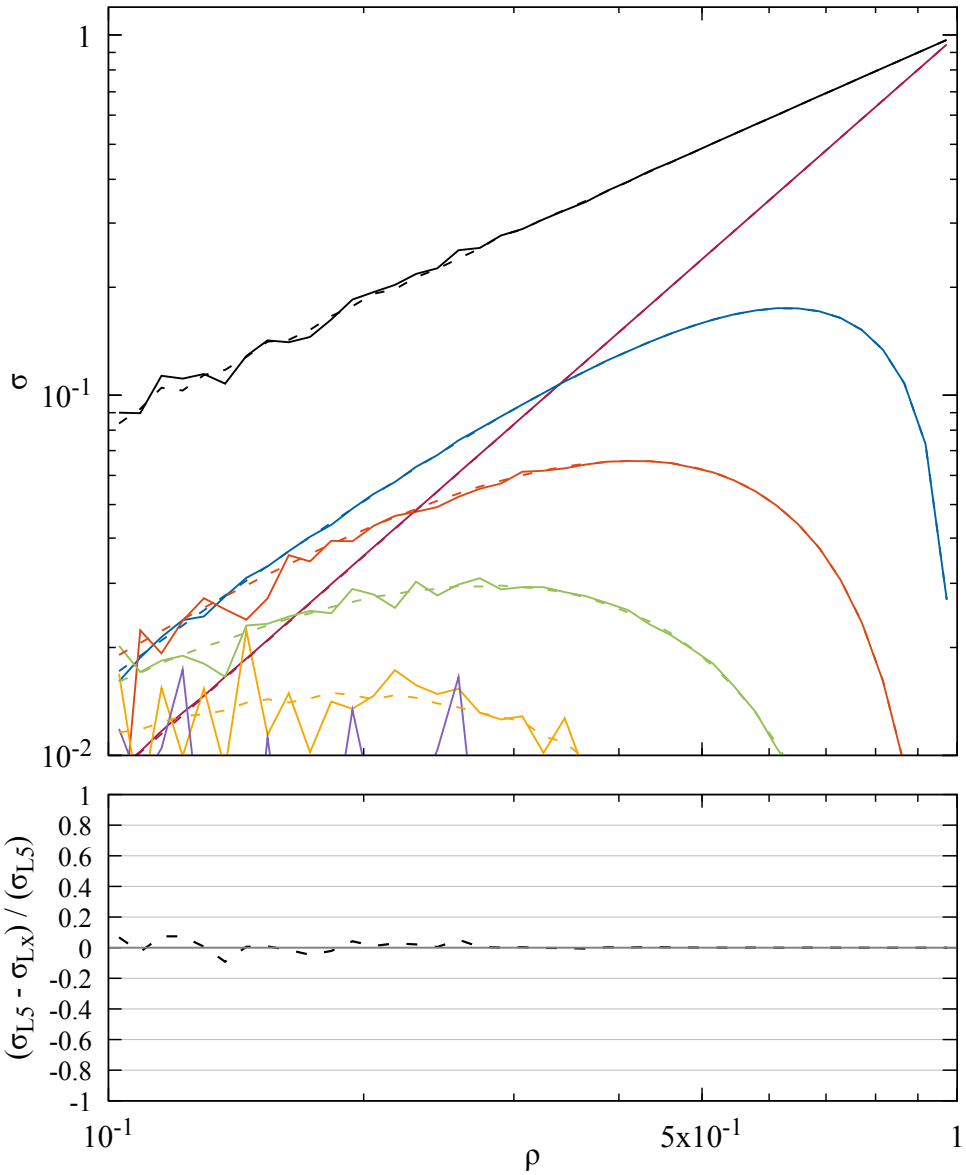


Figure 4.52: The veto cross-section for the tu-interference contribution to $gg \to gg$. Solid: Full colour, Short-dashed: Leading colour.

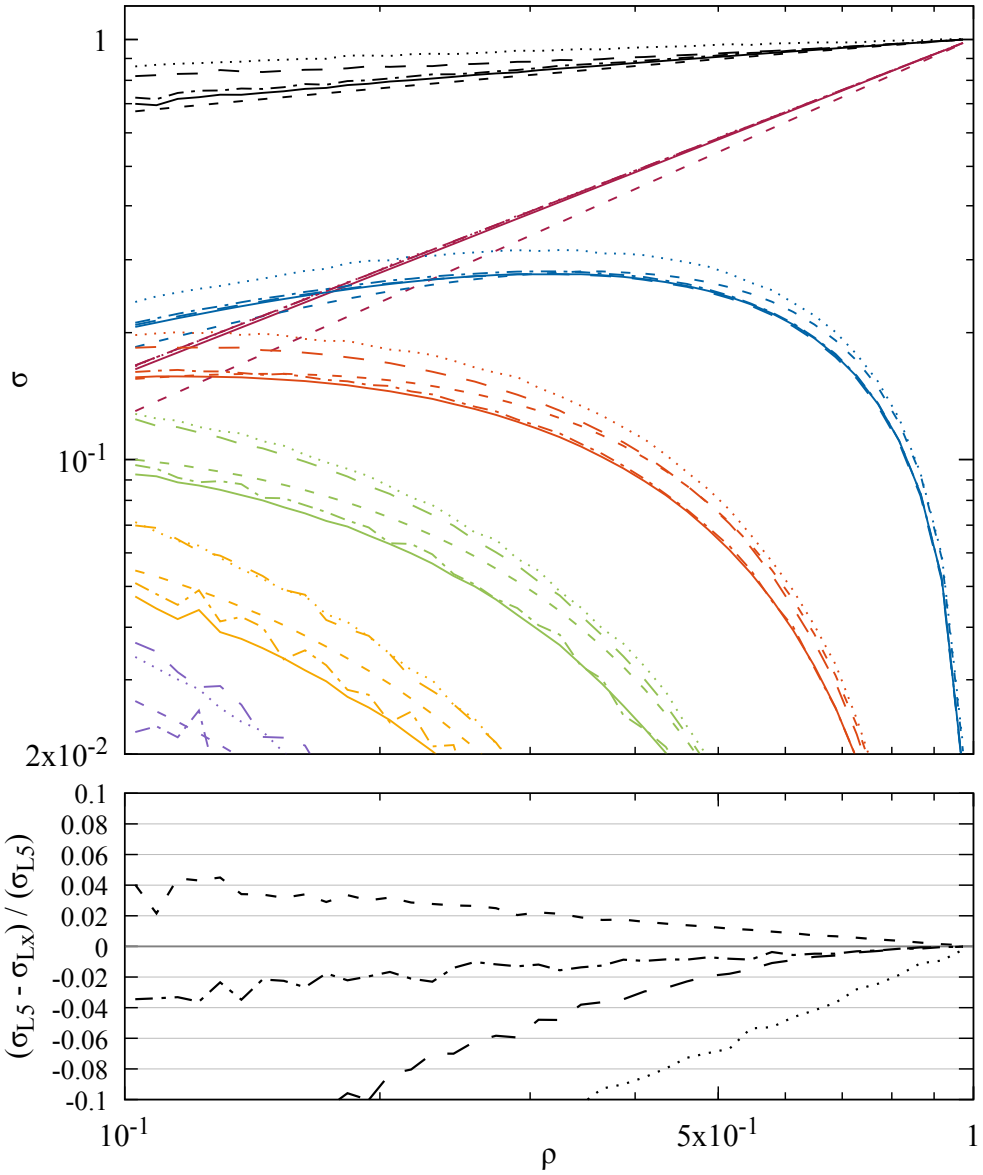


Figure 4.53: The veto cross-section for the $|01\rangle\langle 01|$ contribution to $ZZ \to q\bar{q}q\bar{q}$. Solid: Full colour (L5), Dash-dotted: LC′ + FCR (L4), Long-dashed: LC′ + LCR + singlets (L3), Dotted: LC′ + LCR (L2), Short-dashed: strict LC (L1).

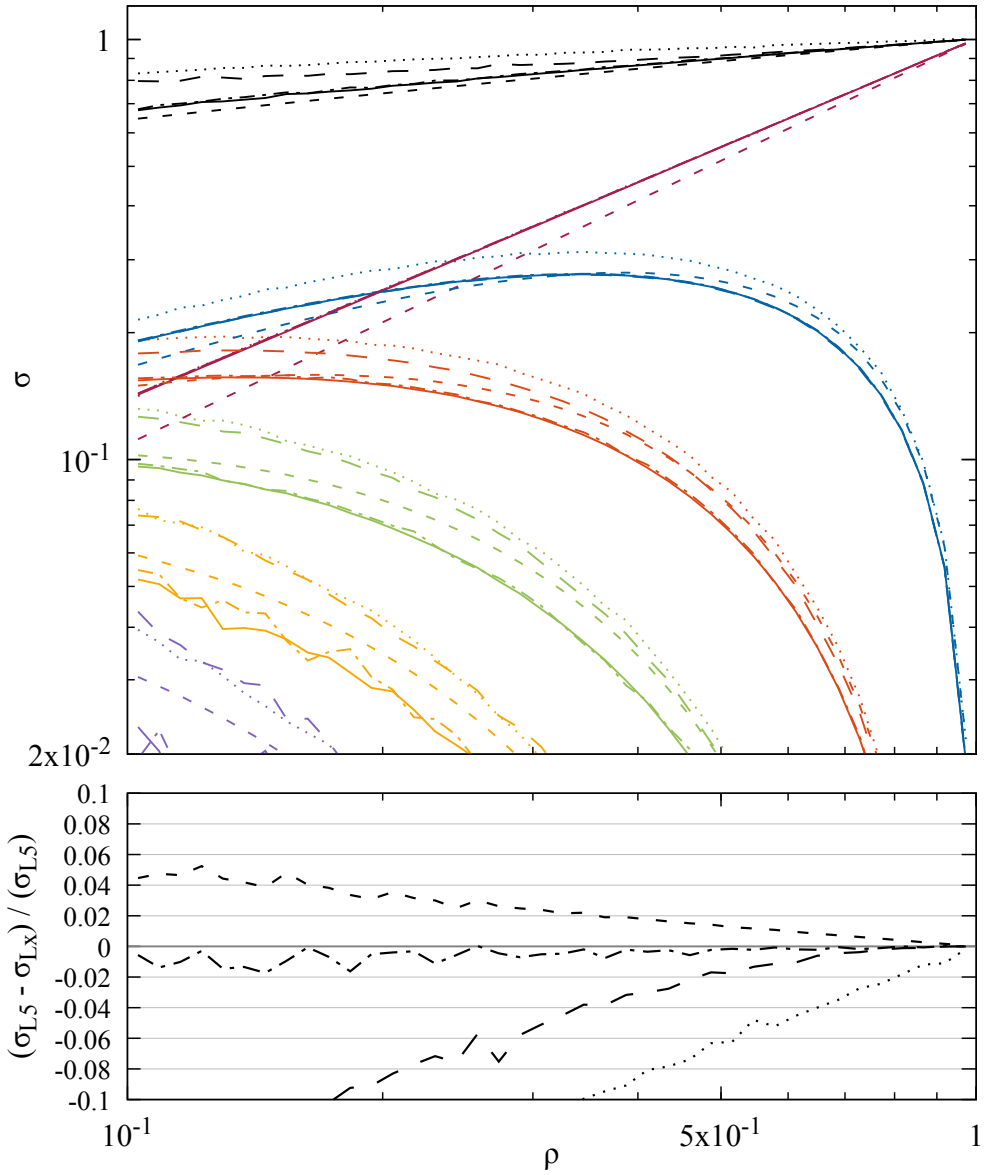


Figure 4.54: The veto cross-section for the $|10\rangle\langle 10|$ contribution to $ZZ \to q\bar{q}q\bar{q}$. Solid: Full colour (L5), Dash-dotted: LC′ + FCR (L4), Long-dashed: LC′ + LCR + singlets (L3), Dotted: LC′ + LCR (L2), Short-dashed: strict LC (L1).

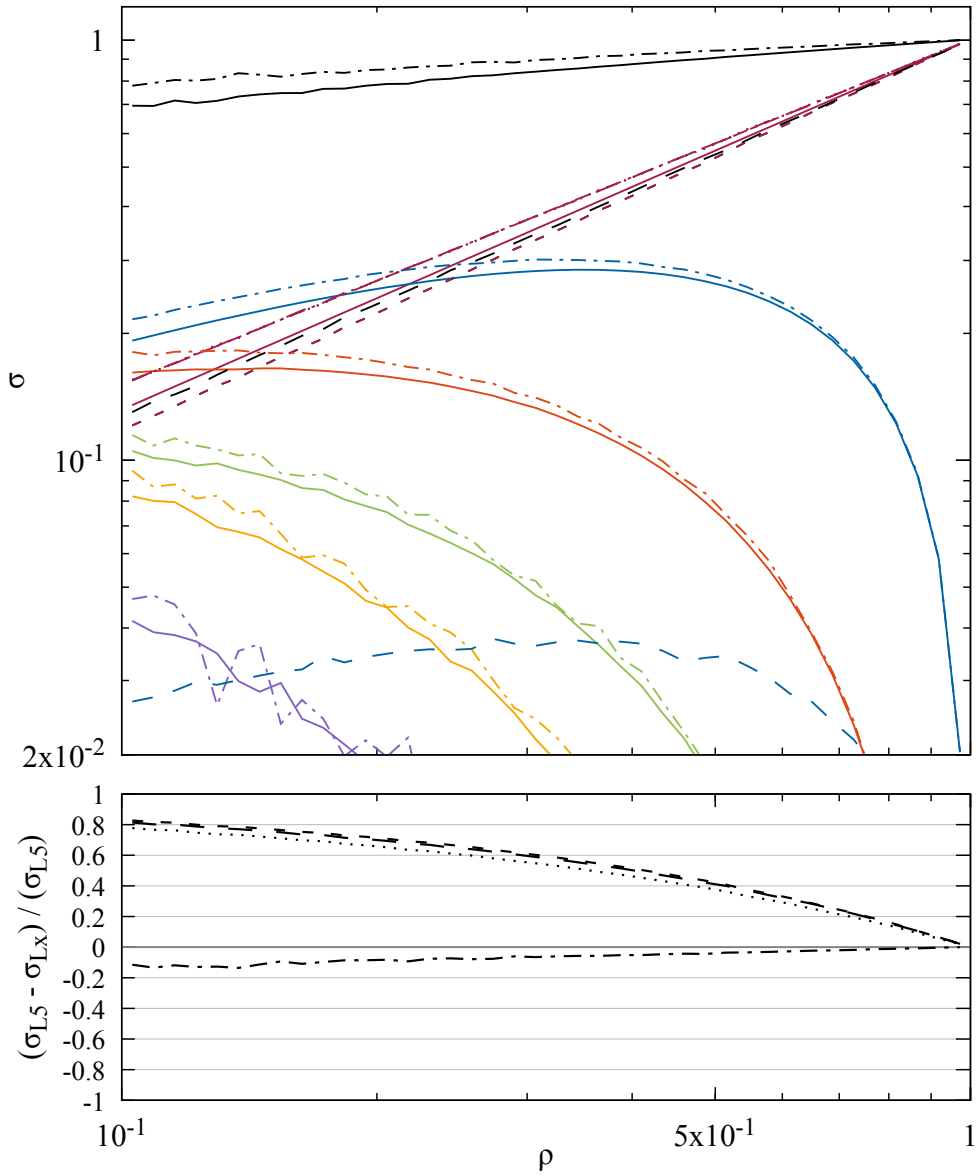


Figure 4.55: The veto cross-section for the $|01\rangle\langle 10|$ (interference) contribution to $ZZ \to q\bar{q}q\bar{q}$. Solid: Full colour (L5), Dash-dotted: LC′ + FCR (L4), Long-dashed: LC′ + LCR + singlets (L3), Dotted: LC′ + LCR (L2), Short-dashed: strict LC (L1). The L3, 1 emission cuve (blue, long-dashed) is negative and the absolute value is plotted.

Chapter 5

Event generator and differential observables

In the previous chapter we investigated in detail the effect of subleading colour in the gaps-between-jets observable, by comparing different approximations to the full-colour result for the cumulative observable, as a function of the veto scale ρ. Now we proceed by employing the event generator mode of `CVolver`, which was described in detail in Section 3.3.2, to analyse the wide-angle soft radiation differentially. This chapter is structured as follows: In Section 5.1 we compare the gaps-between-jets cross-section between dedicated mode and event generator mode for $q\bar{q} \to q\bar{q}$, and we also perform unitarity tests. In Section 5.2 we define the differential observables we will be using for the rest of the chapter. In Sections 5.3 through 5.6 we present differential results for different kinematic configurations and discuss the subleading colour effects.

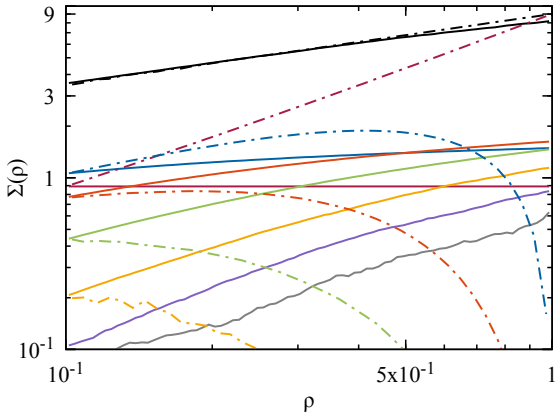
5.1 Dedicated mode agreement and unitarity check

As a cross-check of the event generator mode of `CVolver`, we compare the cumulative gaps-between-jets cross-section with the results from dedicated mode. This is presented, broken down by multiplicity, in Fig. 5.1. The process is $q\bar{q} \to q\bar{q}$ in the fully-symmetric final state kinematics, defined in Table 5.1, with a gap in the region $3\pi/4 < \cos\theta < \pi/4$. It is equivalent to the "symmetric configuration" we considered in Chapter 4. The colours used to represent each multiplicity will remain the same for the rest of this chapter.

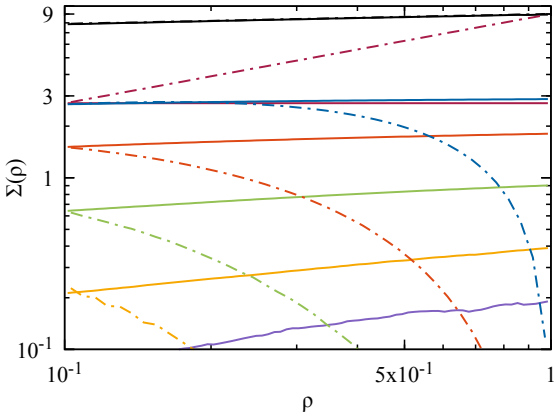
For the $|10\rangle\langle 10|$ and $|10\rangle\langle 01|$ contributions we observe perfect agreement between the two modes. This is especially non-trivial as the multiplicities build the cross-section differently: in dedicated mode all emissions are out of gap, while in event generator mode emissions can be in any region, and contribute cumulatively above the scale of the highest-energy emission in the gap. The event generator data was evolved down to an infrared scale $\mu = 0.1$, and to contribute to the cross-section at $\rho = 0.1$ all emissions must have been out of the gap. Therefore, for the lowest ρ bin the event generator mode and dedicated mode are equivalent.

We do not see agreement between the two modes at high ρ for the $|01\rangle\langle 01|$ evolution. This colour configuration radiates more into the gap region than $|10\rangle\langle 01|$ and $|10\rangle\langle 10|$ (see Section 5.3), and at high ρ the contributions from higher multiplicities, for example 5 or 6, come at the same order as the contributions from 3 and 4 emissions in $|10\rangle\langle 10|$. We can estimate based on the magnitude of the disagreement and the contributions from each multiplicity that the 7 and 8 multiplicity curves are required to capture correctly the high ρ cross-section in $|01\rangle\langle 01|$.

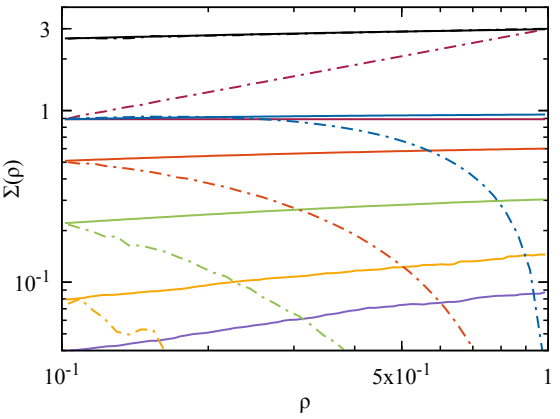
We can also compute the fully-inclusive (no gap) cross-section in order to test unitarity.



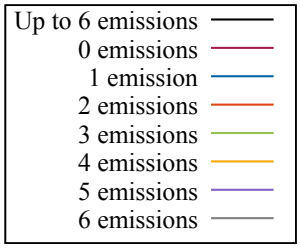
(a) The $|01\rangle\langle 01|$ contribution.



(b) The $|10\rangle\langle 10|$ contribution.



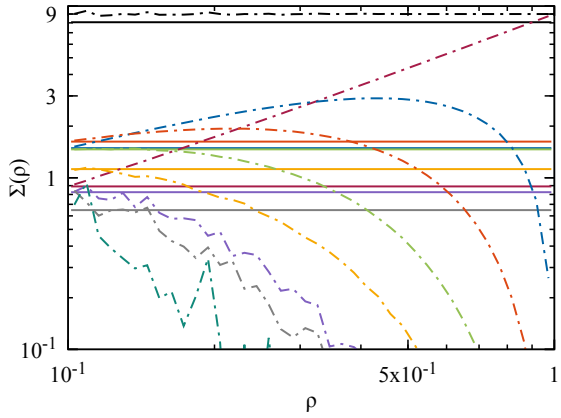
(c) The $|10\rangle\langle 01|$ (interference) contribution.



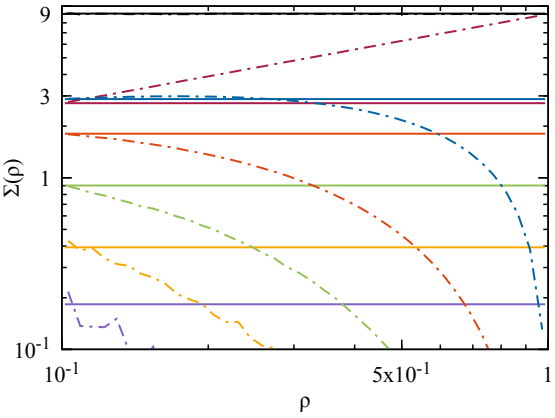
(d) Colours used for each multiplicity.

Figure 5.1: Comparison of the event generator and dedicated modes of `CVolver`, for the $q\bar{q} \to q\bar{q}$ process. The solid curves are the full colour (L5) event generator evolution and the dash-dotted curves are the full colour (L5) dedicated mode evolution. Fig. 5.1d shows the multiplicity legend used for the rest of this chapter.

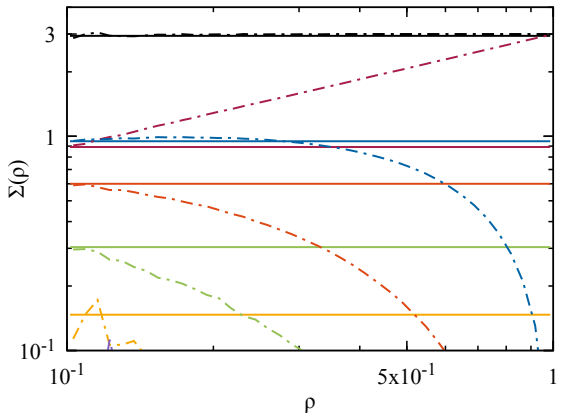
This is presented for both modes in Fig. 5.2. As before, we observe good agreement for the $|10\rangle\langle 01|$ and $|10\rangle\langle 10|$ contributions. The total cross-section remains constant as all virtual and real contributions cancel. This is a highly non-trivial check as the evolution is performed at full-colour. On the other hand, $|01\rangle\langle 01|$ fails the cross-check, for the same reason it failed in Fig. 5.1. We require at least the 7 emissions and 8 emissions curves to restore unitarity. In dedicated mode, 7 emissions was also needed (dash-dotted blue-green curve, underneath dash-dotted grey) to maintain unitarity close to $\rho = 0.1$.



(a) The $|01\rangle\langle 01|$ contribution.



(b) The $|10\rangle\langle 10|$ contribution.



(c) The $|10\rangle\langle 01|$ (interference) contribution.

Figure 5.2: Comparison of the event generator and dedicated modes of `CVolver`, for the $q\bar{q} \to q\bar{q}$ process, broken down by multiplicity and without a gap region, testing unitarity. The solid curves are the full colour (L5) event generator evolution, and the dash-dotted curves are the full colour (L5) dedicated mode evolution.

5.2 Differential observables

The goal is to analyse subleading colour effects in different angular regions of the soft radiation. For this purpose we define the triple-differential cross-section for soft radiation:

$$\frac{\mathrm{d}^3\Sigma}{\mathrm{d}\Omega \mathrm{d}\rho} = \sum_n \int \mathrm{d}\Sigma_n \delta\left(\rho - E_i\right) \delta\left(\Omega - \Omega_i\right) \tag{5.1}$$

where i is the soft emission of highest energy in a gap region Ω_{gap}. Therefore, only events with at least one emission in the gap region will contribute to the differential cross-section. If we integrate this observable over Ω_{gap} and ρ, and include the contributions from events

	Fully-symmetric final state		ϕ-symmetric final state		Fully-asymmetric final state		$0 \to 4$ Fully asymmetric	
	θ	ϕ	θ	ϕ	θ	ϕ	θ	ϕ
q_1	0	0	0	0	0	0	0	0
$\bar{q}_2$	π	0	π	0	π	0	$2\pi/3$	$\pi/12$
q_3	$\pi/6$	0	0.037	0	$\pi/6$	0	$11\pi/12$	0
$\bar{q}_4$	$5\pi/6$	π	2.873	π	$\pi/3$	$19\pi/12$	$3\pi/4$	$\pi/4$

Table 5.1: The directions of the quarks and anti-quarks for each of the kinematic configurations considered.

with no emissions, or all emissions out of gap, the cumulative gaps-between-jets observable presented in Fig. 5.1 is reproduced. This has been checked numerically. We can integrate the triple-differential cross-section to generate a number of different observables:

$$\frac{\mathrm{d}^2\Sigma(r)}{\mathrm{d}\Omega} = \int_0^r \mathrm{d}\rho \frac{\mathrm{d}^3\Sigma}{\mathrm{d}\Omega \mathrm{d}\rho}, \tag{5.2}$$

where we have integrated over ρ to obtain the differential cross-section over solid angle. The upper bound r on the ρ integral vetoes all gluon radiation into the gap region with energy above r. We can also obtain the differential radiation with respect to the polar angle:

$$\frac{\mathrm{d}\Sigma(r)}{\mathrm{d}\cos\theta} = \int_0^{2\pi} \int_0^r \mathrm{d}\rho \, \mathrm{d}\phi \frac{\mathrm{d}^3\Sigma}{\mathrm{d}\Omega \mathrm{d}\rho}, \tag{5.3}$$

and the azimuth:

$$\frac{\mathrm{d}\Sigma(r)}{\mathrm{d}\phi} = \int_0^1 \int_0^r \mathrm{d}\rho \, \mathrm{d}(\cos\theta) \frac{\mathrm{d}^3\sigma}{\mathrm{d}\Omega \mathrm{d}\rho}. \tag{5.4}$$

The last differential cross-section we consider is with respect to the rapidity between jet pairs i and j:

$$\frac{\mathrm{d}\Sigma}{\mathrm{d}y_{ij}} = \int \mathrm{d}\Omega \int_0^r \mathrm{d}\rho \, \delta\left(y_{ij} - \frac{1}{2}\ln\frac{n_i \cdot n(\Omega)}{n_j \cdot n(\Omega)}\right) \frac{\mathrm{d}^3\Sigma}{\mathrm{d}\Omega \mathrm{d}\rho}, \tag{5.5}$$

which is equivalent to boosting the radiation to the frame where i and j are back to back, and evaluating the differential cross-section with respect to rapidity in that frame. The

direction of n_i corresponds to $y_{ij} \to -\infty$, and the direction of n_j corresponds to $y_{ij} \to +\infty$. As a final note on the observables, we will not normalise any of them according to their hard process normalisations, as opposed to the results shown in the previous chapter.

In Section 5.1 we observed that we need higher multiplicity contributions to accurately evaluate the gaps-between-jets observable at high ρ for specific colour states. To circumvent this problem, the analysis of the differential observables (defined in Eqs. (5.2), (5.3), (5.4), and (5.5)) presented in the following sections is performed with $r = 0.3$. This vetoes all radiation above $\rho = 0.3$, and therefore enhances the contribution from low multiplicities, which have a higher probability of surviving without emitting into the gap. This can be seen looking at the region $0.1 < \rho < 0.3$ in Fig. 5.1: for $|10\rangle\langle 01|$ and $|10\rangle\langle 10|$, which have less probability of emitting into the gap, the contribution from each multiplicity remains at the same order of magnitude throughout ρ. On the other hand, for $|01\rangle\langle 01|$, where there is a higher probability of emitting into the gap, the contributions from high multiplicities are suppressed for low values of ρ.

We introduce the differential observables by presenting the soft radiation plotted in a 2D histogram over the angular variables $\cos\theta$ and ϕ, i.e., the differential cross-section defined in Eq. (5.2). We only focus on events with exactly three emissions, in or out of gap. We did not generate more statistics for this process than the others, and these plots are indicative that we are well within reach of analysing the full colour soft radiation differentially. The three contributions to the $ZZ \to q\bar{q}q\bar{q}$ process are presented in Fig. 5.4. The locations of the hard jets are marked on the plot, with colours matching Fig. 5.3, and in this particular case all four jets are outgoing. The kinematics are in the final column of Table 5.1, and the gap is defined as the region outside four cones centred around each of the hard jets, with opening angle $\arccos(0.95)$. The out-of-gap region is indicated by the black areas surrounding the hard jets, where no radiation occurs. A noticeable feature is the strong radiation between q_1 (pink line at $\cos\theta = 1$) and $\bar{q}_2$ (light blue cross) in the $|01\rangle\langle 01|$ plot, which is much weaker for the $|10\rangle\langle 10|$ contribution. In the latter, q_1 is colour connected with $\bar{q}_4$ (red cross), which shifts the radiation slightly higher in ϕ, and is weaker likely due to the wider angle. The radiation of the $|10\rangle\langle 01|$ contribution is an entirely subleading colour effect, because the interference cannot emit at leading colour (L1).

5.3 Fully-symmetric final state, $q\bar{q} \to q\bar{q}$

We continue the analysis by considering the differential observables for the simplest kinematics: the fully-symmetric final state configuration, as defined in Table 5.1. The gap

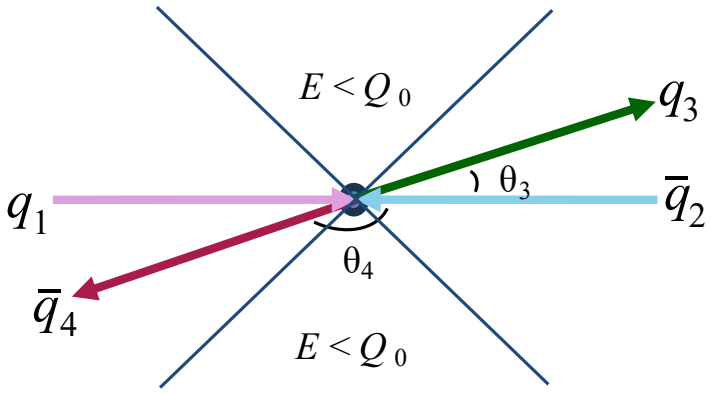
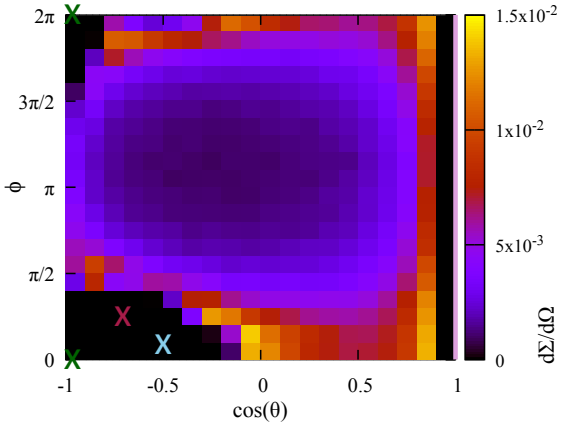


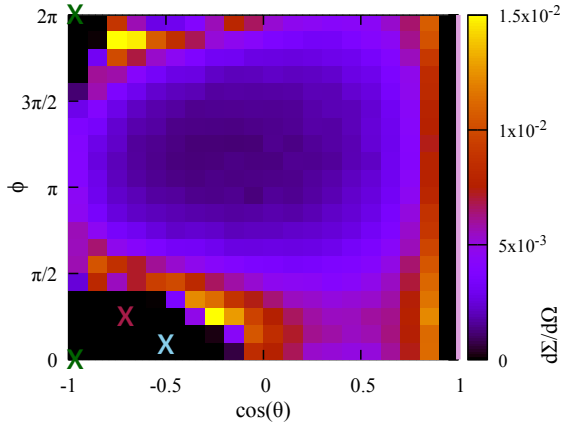
Figure 5.3: Possible kinematic configuration of the four hard legs, with a veto on the energy in a gap region. The angles of the different configurations considered are given in Table 5.1, and the gap definitions for each configuration are given in the text. The colours of the four legs (pink q_1, light blue $\bar{q}_2$, dark green q_3, red $\bar{q}_4$) are used in the figures throughout this chapter to indicate the location of the hard jets. For $2 \to 2$ processes, the colour state $|01\rangle$ corresponds to pink-light blue colour connected, and dark green-red connected. The colour state $|10\rangle$ corresponds to pink-dark green connected, and light blue-red connected. For $0 \to 4$ processes, the colour state $|01\rangle$ corresponds to pink-light blue colour connected, and dark green-red connected. The colour state $|10\rangle$ corresponds to pink-red connected, and light blue-dark green connected.

is defined as the central region $-0.8 < \cos(\theta) < 0.8$. For each observable we plot the $|01\rangle\langle 01|$ and $|10\rangle\langle 10|$ contributions side by side in the top row, and the t-channel and s-channel contributions in the bottom row. The $|01\rangle\langle 01|$ configuration corresponds to the incoming particles being colour connected and the outgoing particles being colour connected (in Fig. 5.3, q_1 connected to q_2 and q_3 connected to q_4). The $|10\rangle\langle 10|$ configuration corresponds to the incoming quark connected to the outgoing quark and equivalently for the antiquarks (q_1 connected to q_3 and q_2 connected to q_4).

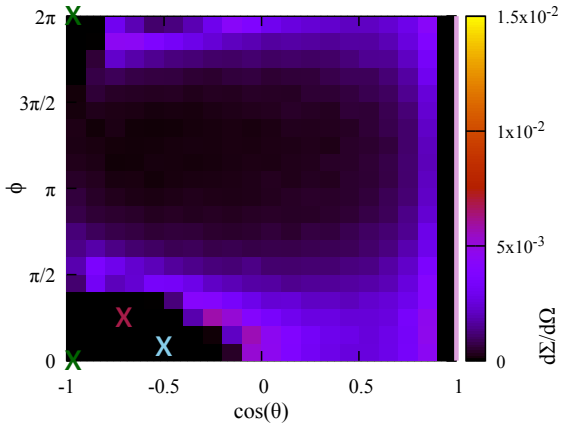
For each contribution we show the full colour (L5) evolution and the strict LC (L1) evolution, and for the s- and t- channel gluon exchange contributions we start the L1 evolution from the leading-colour part of the hard-scatter matrix. The L1, LCH approximation was chosen as it is the best performing approximation to the full-colour gaps-between-jets observable: we found that L1,LCH showed very good agreement with the full colour L5 result for t-channel gluon exchange in every $2 \to 2$ scattering process we considered. By analysing the differential soft radiation patterns we aim to stress-test the success of this approximation.



(a) The $|01\rangle\langle 01|$ contribution.



(b) The $|10\rangle\langle 10|$ contribution.



(c) The $|10\rangle\langle 01|$ (interference) contribution.

Figure 5.4: The full-colour (L5) 3-emission differential cross-section $\mathrm{d}\Sigma/\mathrm{d}\Omega$ of the contributions to the $ZZ \to q\bar{q}q\bar{q}$ process. The location of the hard jets is marked with the colours used in Fig. 5.3 (the four crosses in $\cos(\theta) < 0$, and the pink line at $\cos(\theta) = 1$). The black bins around the jets indicate the out-of-gap regions.

The $\mathrm{d}\Sigma/\mathrm{d}(\cos\theta)$ distributions are shown in Fig. 5.5. As expected, we see the radiation grows as it gets closer to the hard jets (their location is indicated with vertical lines matching the colours in Fig. 5.3). Radiation hits a minimum in the middle of the gap, the furthest point from all jets. The $|01\rangle\langle 01|$ distribution is less steep than $|10\rangle\langle 10|$, which follows from the colour flows: in the former configuration, the colour-connected pairs are back-to-back (light blue is connected to pink, and red is connected to dark green), which causes them to emit in every direction, including the gap region. In contrast, $|10\rangle\langle 10|$ has the connected pairs at angles of $\pi/6$, aiming towards the out-of-gap region (light blue is connected to red, and dark green is connected to pink), which leads to most of the radiation being emitted out of gap, and results in a much steeper soft radiation pattern.

Since L1,LCH and L5 have different normalisations, we should focus on the shape of

the residuals. We see that the residual for the $|10\rangle\langle 10|$ contribution is approximately flat. This is consistent with Fig. 4.12, where the L1 result is only 1-2% away from the L5, and so we do not expect large effects between full colour and leading colour. On the other hand, the residual for the $|01\rangle\langle 01|$ shows a 5-10% effect between the edges of the gap and the middle, although we have some fluctuations. It is also possible to see this effect looking at the 1, 2 or 3 emission curves (blue, orange, light green respectively). It is caused by swaps to the $|10\rangle\langle 10|$ configuration, which radiates more into the out-of-gap region, therefore enhancing the differential cross-section near the gap edges. These results are consistent with the veto cross-section Fig. 4.11, where there is a 20% difference between L1 and L5 at $\rho = 1$. That said, we will have a more clear example (with smaller fluctuations) for the same distribution when we consider the ϕ-symmetric kinematic configuration in Section 5.4.

The t-channel contribution has a flat residual, showing the subleading-colour effects discussed previously for the $|01\rangle\langle 01|$ evolution cancel when including the interference contribution, shown in Fig. 5.10, and the subleading $|10\rangle\langle 10|$ contribution. Again, this is in agreement with Fig. 4.16, where L1, starting from the LCH, is a good approximation to the full-colour result. We note that the interference evolution cannot be approximated at all at L1, as there would be no radiation, and also that its radiation pattern is different (steeper) from those of the $|01\rangle\langle 01|$ and $|10\rangle\langle 10|$ contributions.

Lastly, the s-channel contribution shows the largest discrepancy between the full colour and leading colour results. The inclusion of the subleading $|01\rangle\langle 01|$ and $|10\rangle\langle 01|$ contributions flatten the full-colour radiation pattern, causing an overall 40% effect between the edges and the middle of the gap.

We move on to the $\mathrm{d}\Sigma/\mathrm{d}\phi$ distribution, in Fig. 5.6. The positions in ϕ of the two outgoing particles are shown with vertical lines on the plot. Again, we see that the the $|01\rangle\langle 01|$ contribution radiates more evenly in ϕ than the $|10\rangle\langle 10|$ contribution, which prioritises radiation along the directions of the two colour-connected pairs. The $|10\rangle\langle 10|$ residual is approximately flat, while the $|01\rangle\langle 01|$ residual shows $\sim 5\%$ effects between the directions of the outgoing particles and the in-between regions. The t-channel residual is approximately flat, although with fluctuations, while the s-channel one shows $\sim 10\%$ effects.

We now consider the $\mathrm{d}\Sigma/\mathrm{d}y_{ij}$ observables. As a reminder, these are equivalent to boosting the radiation pattern to the frame where are i and j are back-to-back, where i is at $y = -\infty$ and j is at $y = \infty$. The gap region gets boosted too, which results in highly-asymmetrical radiation patterns. For example, in Fig. 5.7, the radiation extends further in the region $y_{13} > 0$ than in $y_{13} < 0$, due to q_1 being located in the middle of the out-of-gap

region, while q_3 is closer to the boundary. We note that the outermost bins of the distribution will have strong edge effects, as the gap boundary will not be aligned with the bin boundaries, which was not the case for the $\mathrm{d}\Sigma/\mathrm{d}(\cos\theta)$ and $\mathrm{d}\Sigma/\mathrm{d}\phi$ distributions. We will ignore the behaviour at the edges during the following discussion.

Due to the symmetries of this kinematic configuration, some of the rapidity distributions are equivalent to one another: $\mathrm{d}\Sigma/\mathrm{d}y_{13} \sim \mathrm{d}\Sigma/\mathrm{d}y_{24}$, and $\mathrm{d}\Sigma/\mathrm{d}y_{14} \sim \mathrm{d}\Sigma/\mathrm{d}y_{23}$, so we only need to consider a subset of these. The $\mathrm{d}\Sigma/\mathrm{d}y_{34}$ distribution differs from $\mathrm{d}\Sigma/\mathrm{d}y_{12}$, and therefore from $\mathrm{d}\Sigma/\mathrm{d}(\cos\theta)$, by the rotated gap with respect to the centre-of-mass frame, even though a boost is not required.

All three distributions $\mathrm{d}\Sigma/\mathrm{d}y_{13}$ (Fig. 5.7), $\mathrm{d}\Sigma/\mathrm{d}y_{14}$ (Fig. 5.8), and $\mathrm{d}\Sigma/\mathrm{d}y_{34}$ (Fig. 5.9) follow similar trends: the shapes of the $|01\rangle\langle 01|$ and $|10\rangle\langle 10|$ contributions can be broadly understood in terms of which hard legs are colour connected. The $|10\rangle\langle 10|$ residual is approximately flat, the $|01\rangle\langle 01|$ residual shows $\sim$ 10% effects, and the L1 approximation completely fails for the s-channel gluon exchange. More interestingly, the t-channel residuals are not flat, although the effects are somewhat subtle: in all three plots there are $\sim$ 10% effects between the central rapidity regions and the tails of the distributions. There are also shape differences between the L5 and the L1,LCH radiation distributions. It follows that the exceptional agreement found previously for the t-channel gluon exchange gaps-between-jets cross-section between L5 and L1 does not always hold when considering the differential distributions.

5.4 ϕ-symmetric final state, $q\bar{q} \to q\bar{q}$

We move on the the ϕ-symmetric kinematic configuration, as defined in Table 5.1. The gap is defined as the central region $-0.9 < \cos(\theta) < 0.9$. It is the same as the "asymmetric configuration" in Chapter 4.

The $\mathrm{d}\Sigma/\mathrm{d}(\cos\theta)$ distribution in Fig. 5.11 clearly shows the effect of the small, and asymmetric, scattering angles: the $|10\rangle\langle 10|$ contribution is now highly asymmetric in θ. The q_1 and q_3 opening is much smaller than the q_2 and q_4 opening, suppressing the former dipole, and enhancing radiation around the latter dipole. On the other hand, the $|01\rangle\langle 01|$ contribution, where the colour-connected pairs are almost back-to-back, remains more symmetric, but not completely: full-colour evolution can swap from this colour configuration to the other, which visibly enhances the L5 curves towards $\cos\theta = -1$. This is most visible for the L5, 1 emission curve (solid blue), but is present in all of them.

The $|10\rangle\langle 10|$ evolution emits mostly out-of-gap, and disfavours swapping to a different colour configuration, as doing so would increase radiation into the gap, causing it to get vetoed. This results in a very good agreement between L5 and L1 (as we saw in Fig. 4.19). On the other hand, the $|01\rangle\langle 01|$ contribution prefers swapping over, in order to enhance the radiation out-of-gap, which manifests as large subleading colour effects. This is the same effect we saw for the fully-symmetric kinematics in Fig. 5.5, but made more dramatic.

The s-channel gluon exchange contribution has massive deviations between L5 and L1, as the latter never accesses the configuration that radiates strongly in the gap region. Interestingly, the t-channel L5 result sees the enhancement near $\cos\theta = -1$ of the $|10\rangle\langle 10|$ evolution completely cancel with the interference (Fig. 5.14) and the subleading $|10\rangle\langle 10|$ contributions.

Other differential distributions are shown in Figs. 5.12 and 5.13, and the interferences are shown in Fig. 5.14. We note that all of them show the same characteristics: the $|10\rangle\langle 10|$ contribution has very good agreement between L5 and L1. The $|01\rangle\langle 01|$ contribution shows L5 enhances the radiation in the region near the hard jets. The s-channel gluon exchange completely fails to be described by L1 evolution. Lastly, the t-channel gluon exchange shows the enhancement in $|01\rangle\langle 01|$ cancels with the $|10\rangle\langle 10|$ and interference contributions, resulting in good agreement between L5 and L1.

5.5 Fully-asymmetric final state, $q\bar{q} \to q\bar{q}$

We continue with the fully-asymmetric final-state kinematics, as defined in Table 5.1. The gap is defined as the region outside four cones centred around each of the hard jets, with opening angle $\arccos(0.95)$. In this configuration both final state jets point in the general direction of the incoming q_1, thus not favouring any particular colour configuration. The goal is to test whether this will spoil the agreement between full colour and leading colour for the radiation patterns of the t-channel gluon exchange. We find the evolution of the individual $|01\rangle\langle 01|$ and $|10\rangle\langle 10|$ contributions do not show any remarkable features, likely due to the angles of the connected pairs being of similar magnitudes between one another, so we omit them for brevity. We focus on the s- and t- channel gluon exchange contributions, in Figs. 5.15 and 5.16.

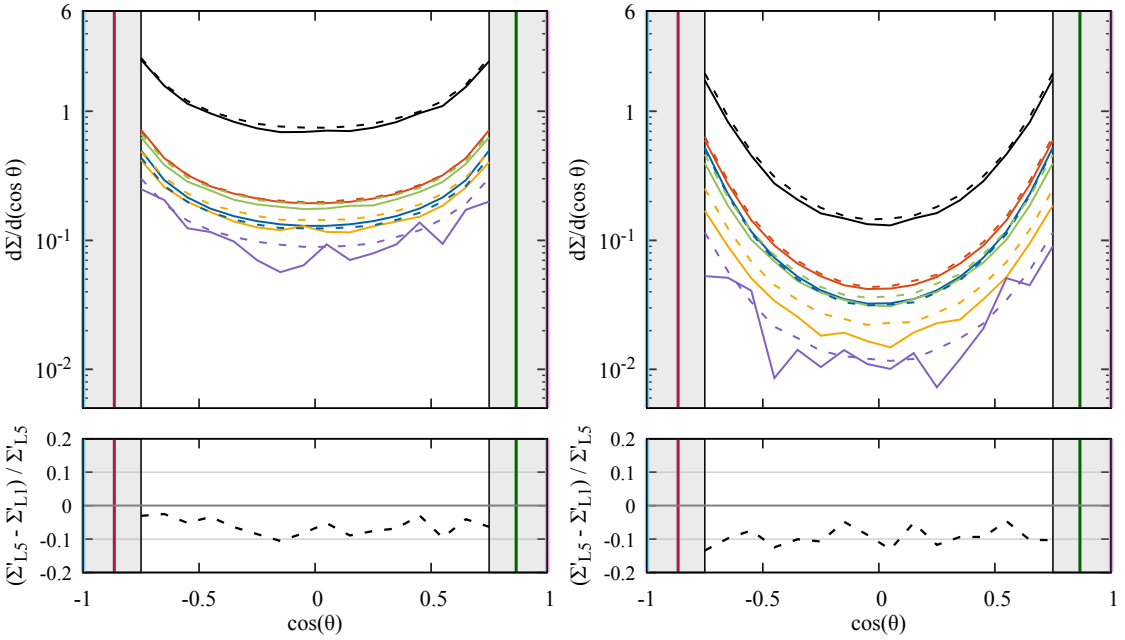
In general, we find that the full-colour (L5) cross-section is enhanced relative to the leading colour (L1) in the region near the three closer jets q_1, q_3, and $\bar{q}_4$ for both gluon exchange channels. It is clearly visible for example in Fig. 5.15. We can also observe significantly different radiation shapes between L5 and L1, which are most clearly visible

looking at individual multiplicities. For example, in Fig. 5.16, t-channel gluon exchange, the 2 emissions orange L5 and L1 curves are different. There are multiple competing colour configurations which radiate in this region, which might explain why full-colour evolution, which can access all competing processes, may enhance radiation.

5.6 Colour singlet production of four jets

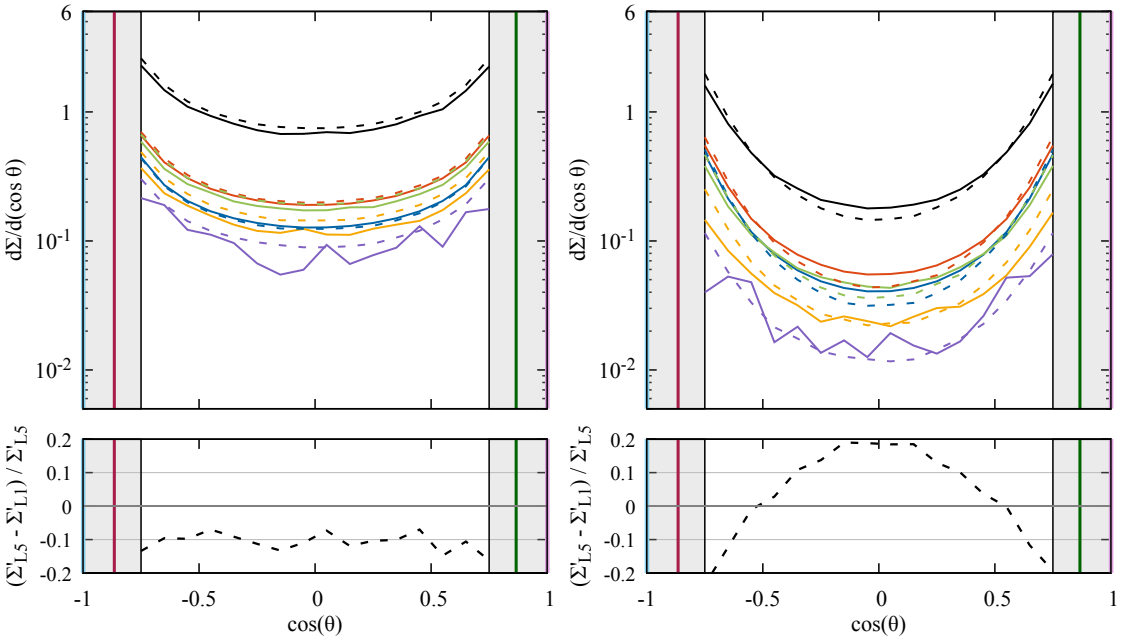
Lastly, we consider the production of four jets in $ZZ \to q\bar{q}q\bar{q}$, with the kinematics as in Table 5.1. The gap is defined as the region outside four cones centred around each of the hard jets, with opening angle $\arccos(0.95)$. This is the same configuration we plotted over a 2D solid angle histogram in Fig. 5.4. We continue using the colours of the four jets in Fig. 5.3, however now q_1 and $\bar{q}_2$ are also outgoing. There are two possible production amplitudes which correspond with a different colour flow each. The first is $|01\rangle$, in which q_1 is connected to $\bar{q}_2$ (pink with light blue) and q_3 is connected to $\bar{q}_4$ (dark green with red). The second is $|10\rangle$, in which q_1 is connected to $\bar{q}_4$ (pink with red) and $\bar{q}_2$ is connected to q_3 (light blue with dark green).

We present the $\cos\theta$ and ϕ differential distributions for the $|01\rangle\langle 01|$, $|10\rangle\langle 10|$ and $|10\rangle\langle 01|$ contributions in Fig. 5.17. We find that there is good agreement between L5 and L1 for the $|01\rangle\langle 01|$ and $|10\rangle\langle 10|$ contributions, up to fluctuations. However, the most interesting feature is that the interference distributions have significantly different radiation patterns compared to the colour-diagonal contributions. The $\mathrm{d}\Sigma/\mathrm{d}\phi$ interference radiation in particular is much steeper in the region away from the hard jets.



(a) The $|01\rangle\langle 01|$ contribution.

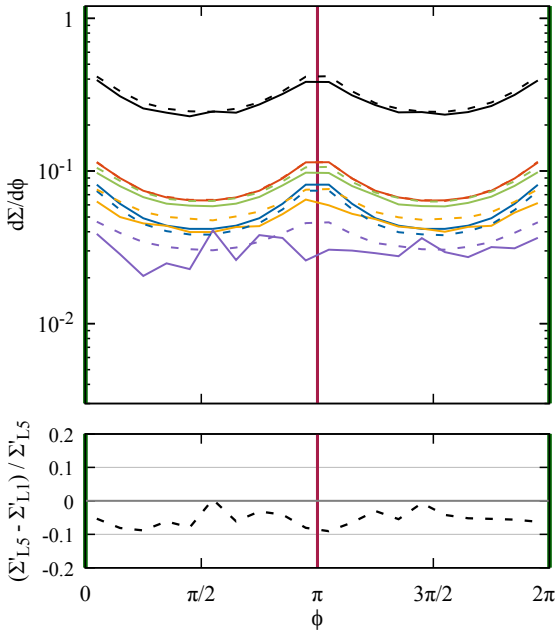
(b) The $|10\rangle\langle 10|$ contribution.



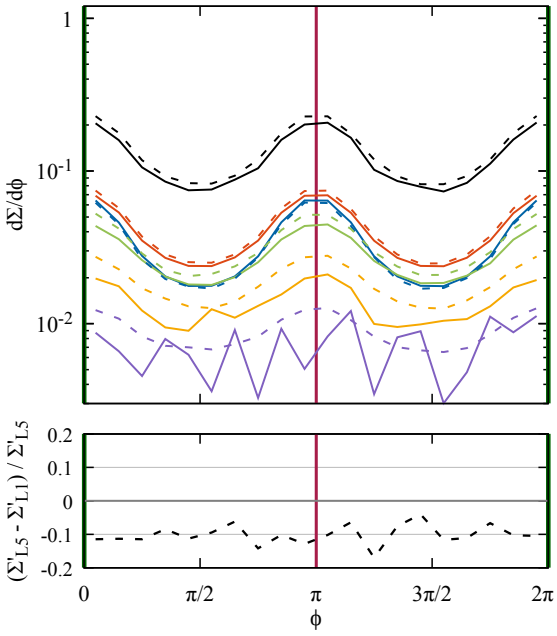
(c) The t-channel gluon exchange contribution.

(d) The s-channel gluon exchange contribution.

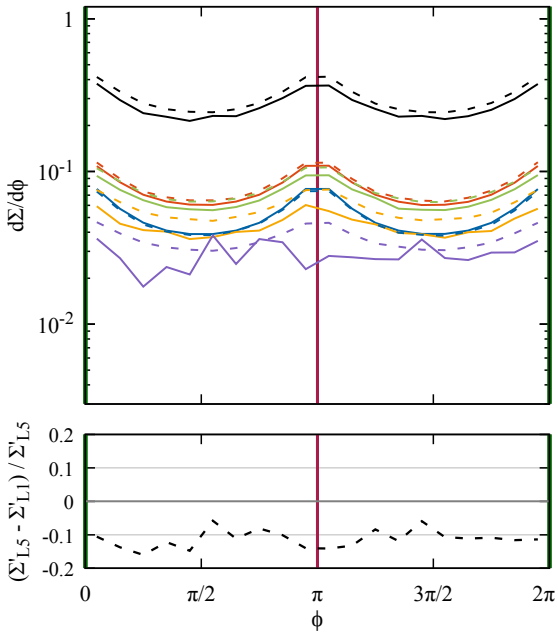
Figure 5.5: The differential cross-section $\mathrm{d}\Sigma/\mathrm{d}(\cos\theta)$ of the different contributions to the $q\bar{q} \to q\bar{q}$ process, in the fully-symmetric final state configuration, and broken down by multiplicity. The solid curves are the full colour (L5) evolution, and the dashed curves are the strict LC (L1) evolution. For the L1 curves we start the evolution using the leading-colour approximation to the hard-scatter matrix. The locations of the hard jets are marked with vertical lines matching the colours used in Fig. 5.3. The shaded vertical bars indicate the out-of-gap regions.



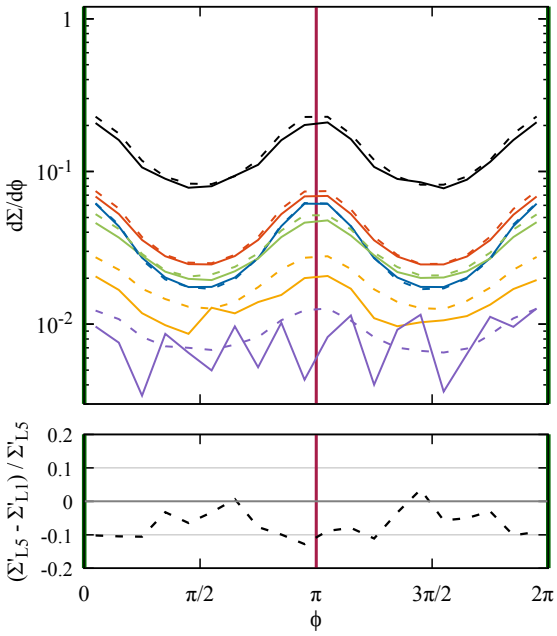
(a) The $|01\rangle\langle 01|$ contribution.



(b) The $|10\rangle\langle 10|$ contribution.

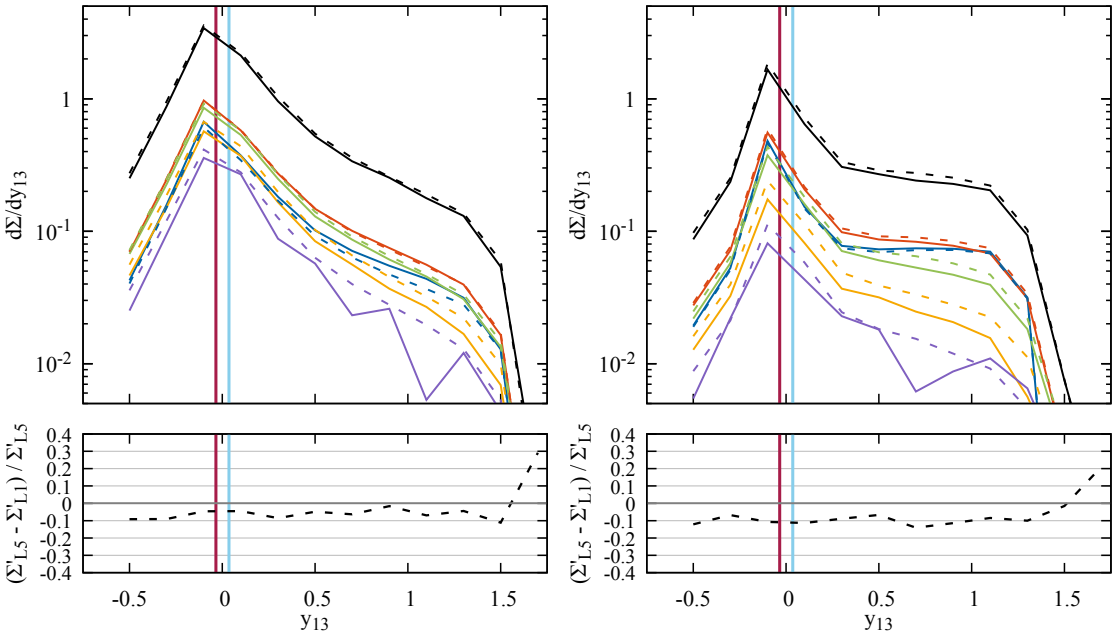


(c) The t-channel gluon exchange contribution.



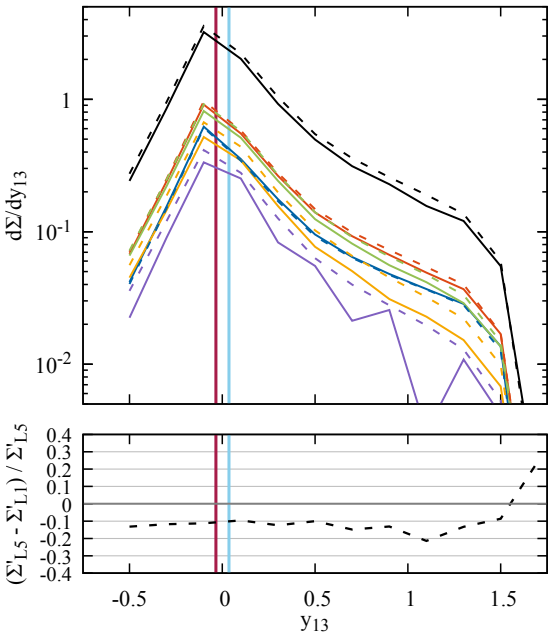
(d) The s-channel gluon exchange contribution.

Figure 5.6: The differential cross-section $\mathrm{d}\Sigma/\mathrm{d}\phi$ of the different contributions to the $q\bar{q} \to q\bar{q}$ process, in the fully-symmetric final state configuration, and broken down by multiplicity. The solid curves are the full colour (L5) evolution, and the dashed curves are the strict LC (L1) evolution. For the L1 curves we start the evolution using the leading-colour approximation to the hard-scatter matrix. The locations of the hard jets are marked with vertical lines matching the colours used in Fig. 5.3.

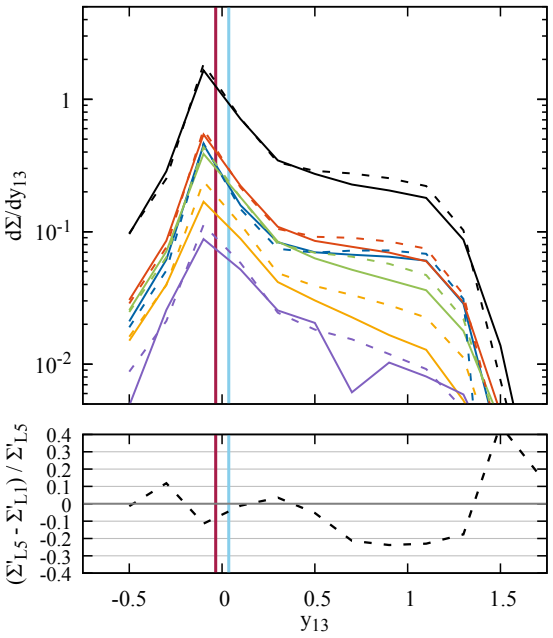


(a) The $|01\rangle\langle 01|$ contribution.

(b) The $|10\rangle\langle 10|$ contribution.

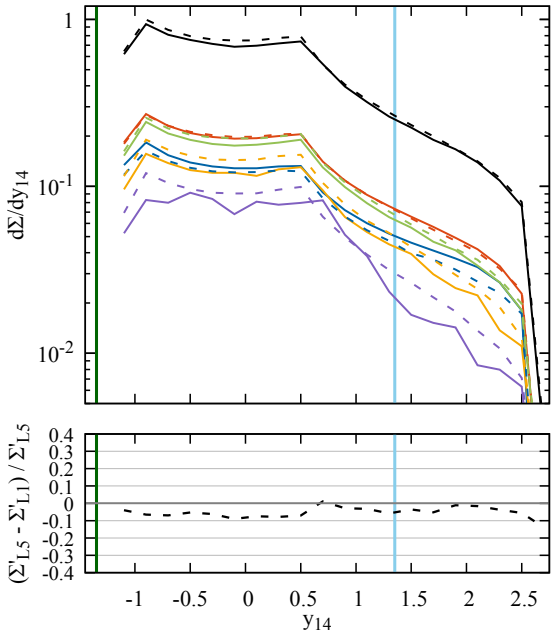


(c) The t-channel gluon exchange contribution.

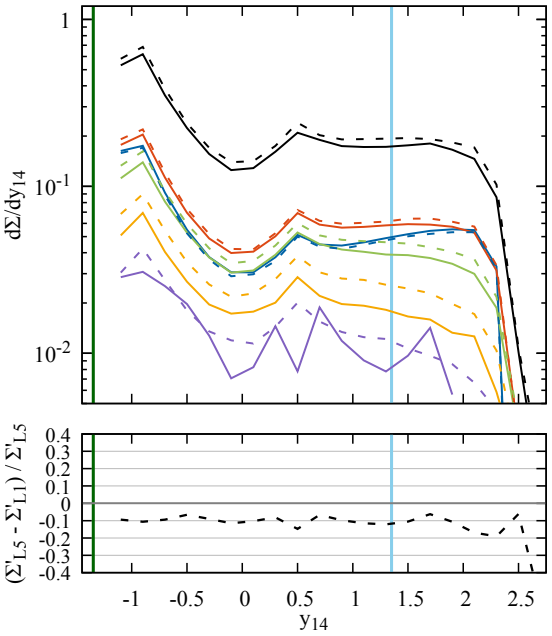


(d) The s-channel gluon exchange contribution.

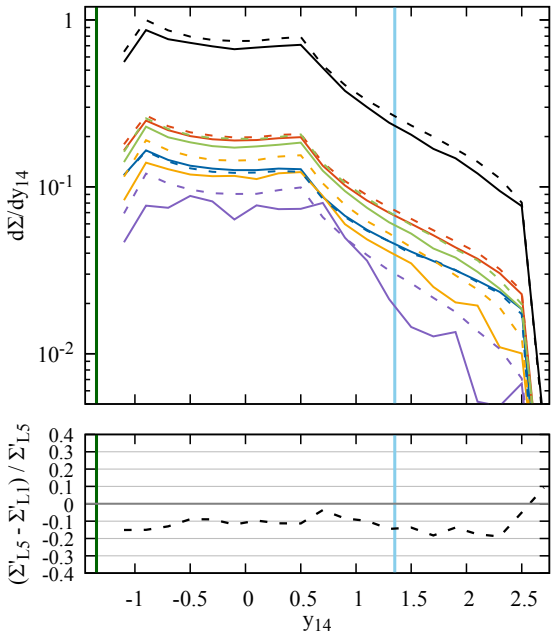
Figure 5.7: The differential cross-section $\mathrm{d}\Sigma/\mathrm{d}y_{13}$ of the different contributions to the $q\bar{q} \to q\bar{q}$ process, in the fully-symmetric final state configuration, and broken down by multiplicity. The solid curves are the full colour (L5) evolution, and the dashed curves are the strict LC (L1) evolution. For the L1 curves we start the evolution using the leading-colour approximation to the hard-scatter matrix. The locations of the hard jets are marked with vertical lines matching the colours used in Fig. 5.3.



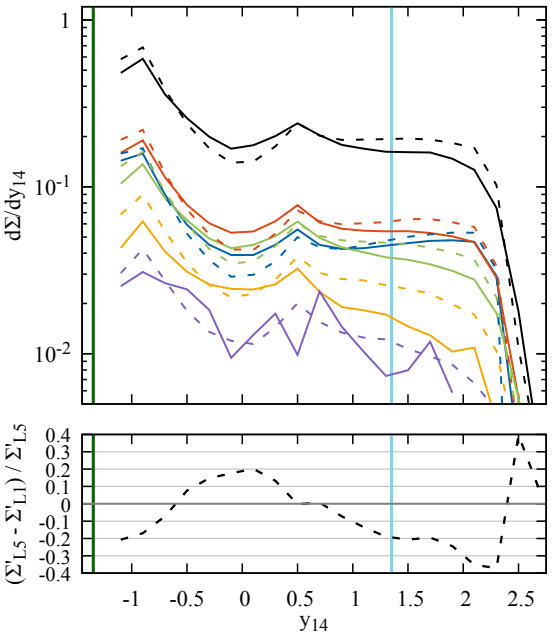
(a) The $|01\rangle\langle 01|$ contribution.



(b) The $|10\rangle\langle 10|$ contribution.

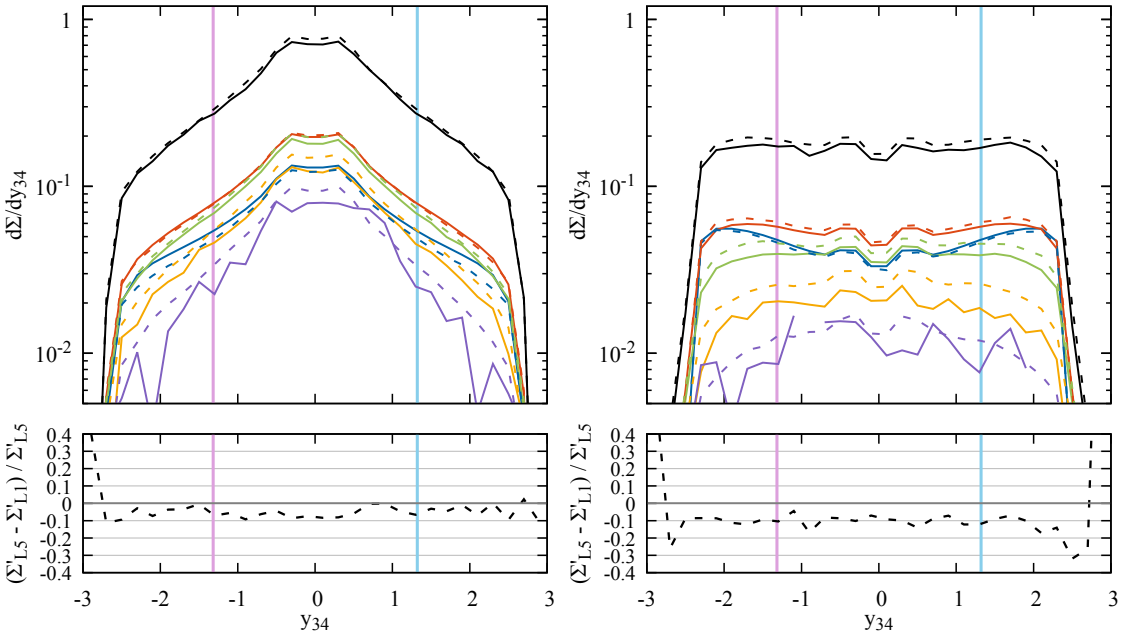


(c) The t-channel gluon exchange contribution.



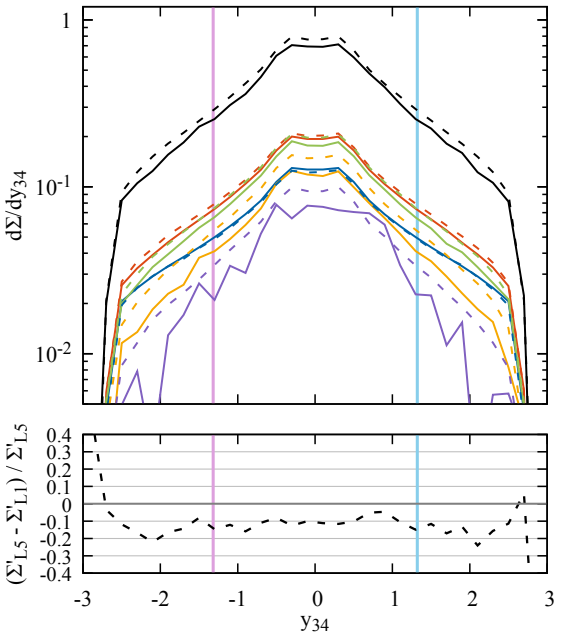
(d) The s-channel gluon exchange contribution.

Figure 5.8: The differential cross-section $\mathrm{d}\Sigma/\mathrm{d}y_{14}$ of the different contributions to the $q\bar{q} \to q\bar{q}$ process, in the fully-symmetric final state configuration, and broken down by multiplicity. The solid curves are the full colour (L5) evolution, and the dashed curves are the strict LC (L1) evolution. For the L1 curves we start the evolution using the leading-colour approximation to the hard-scatter matrix. The locations of the hard jets are marked with vertical lines matching the colours used in Fig. 5.3.

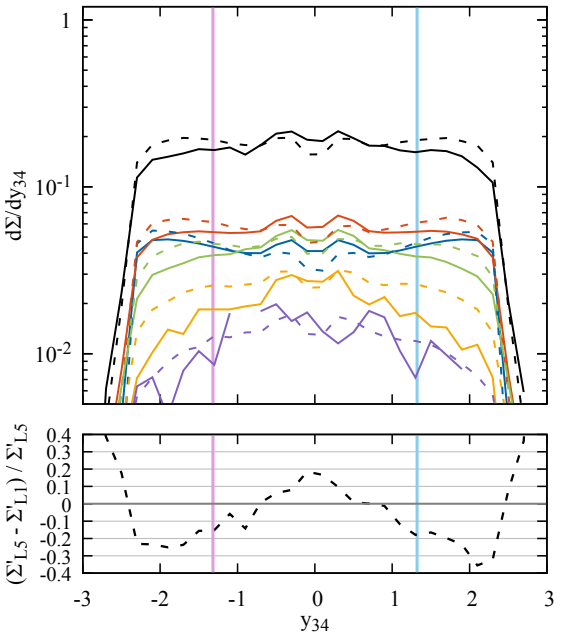


(a) The $|01\rangle\langle 01|$ contribution.

(b) The $|10\rangle\langle 10|$ contribution.



(c) The t-channel gluon exchange contribution.



(d) The s-channel gluon exchange contribution.

Figure 5.9: The differential cross-section $\mathrm{d}\Sigma/\mathrm{d}y_{34}$ of the different contributions to the $q\bar{q} \to q\bar{q}$ process, in the fully-symmetric final state configuration, and broken down by multiplicity. The solid curves are the full colour (L5) evolution, and the dashed curves are the strict LC (L1) evolution. For the L1 curves we start the evolution using the leading-colour approximation to the hard-scatter matrix. The locations of the hard jets are marked with vertical lines matching the colours used in Fig. 5.3.

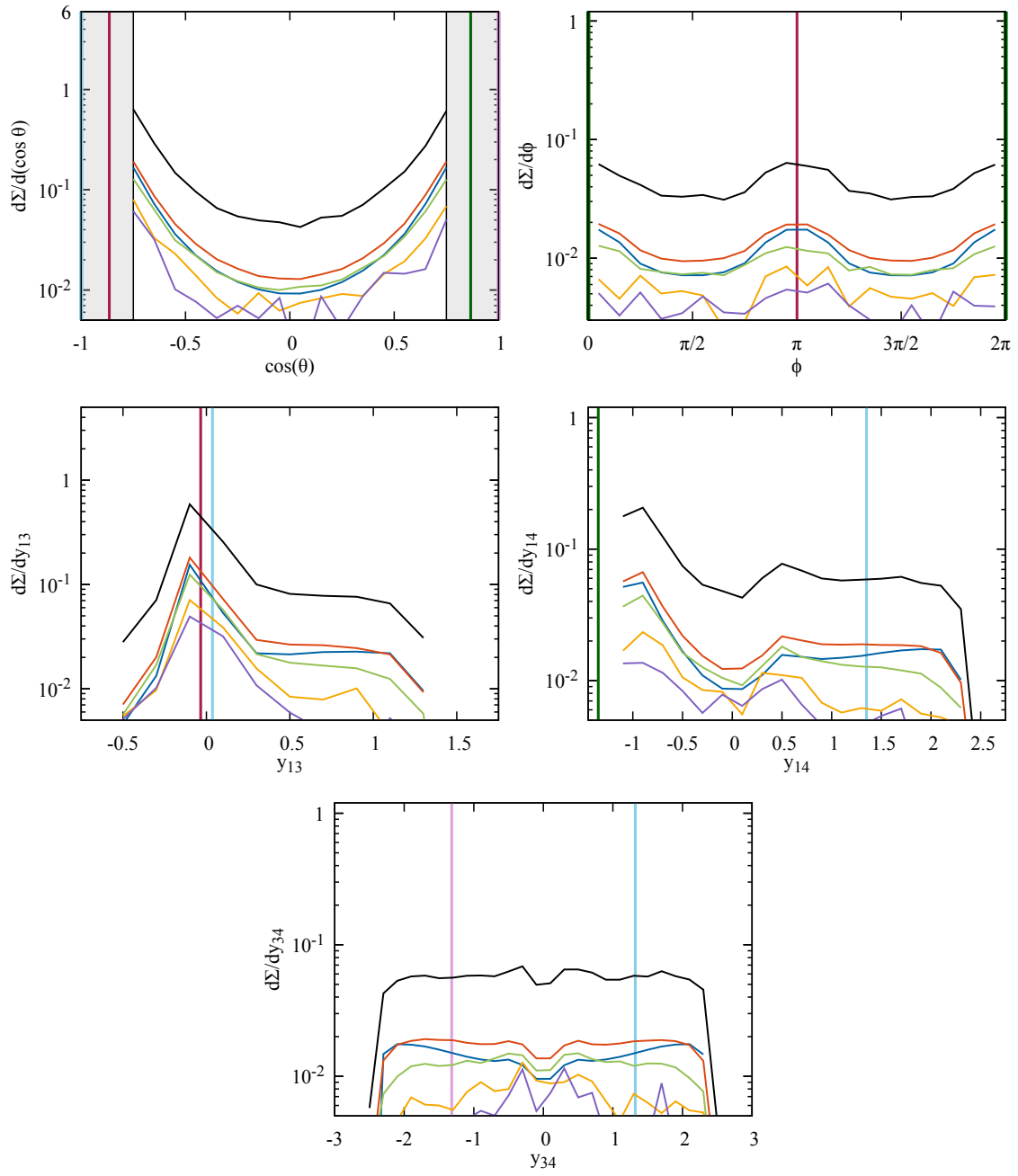
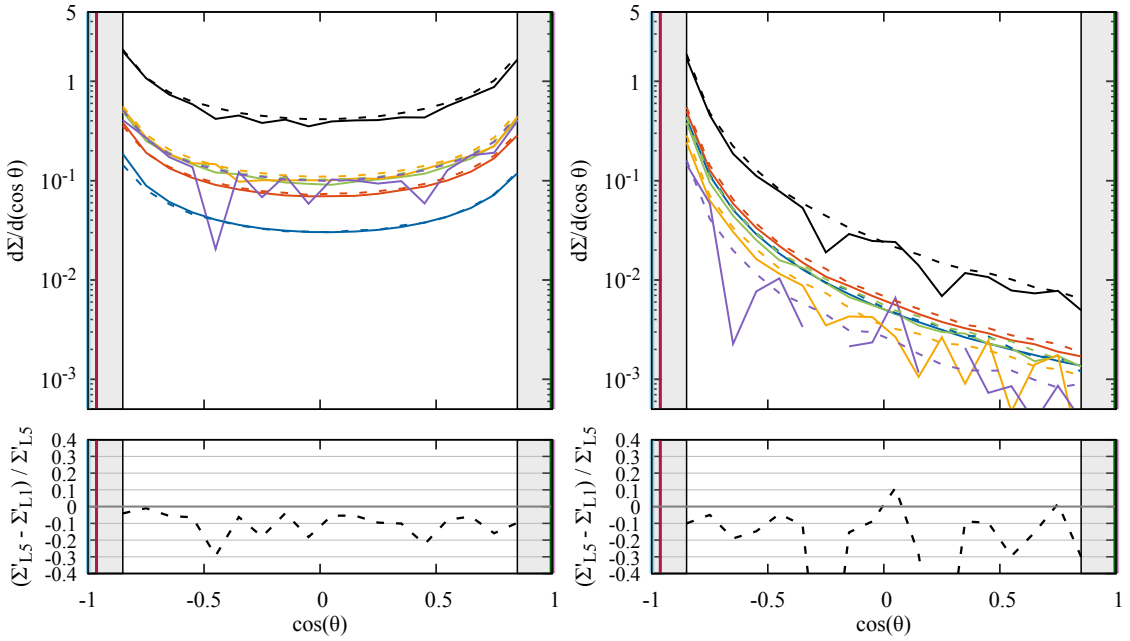
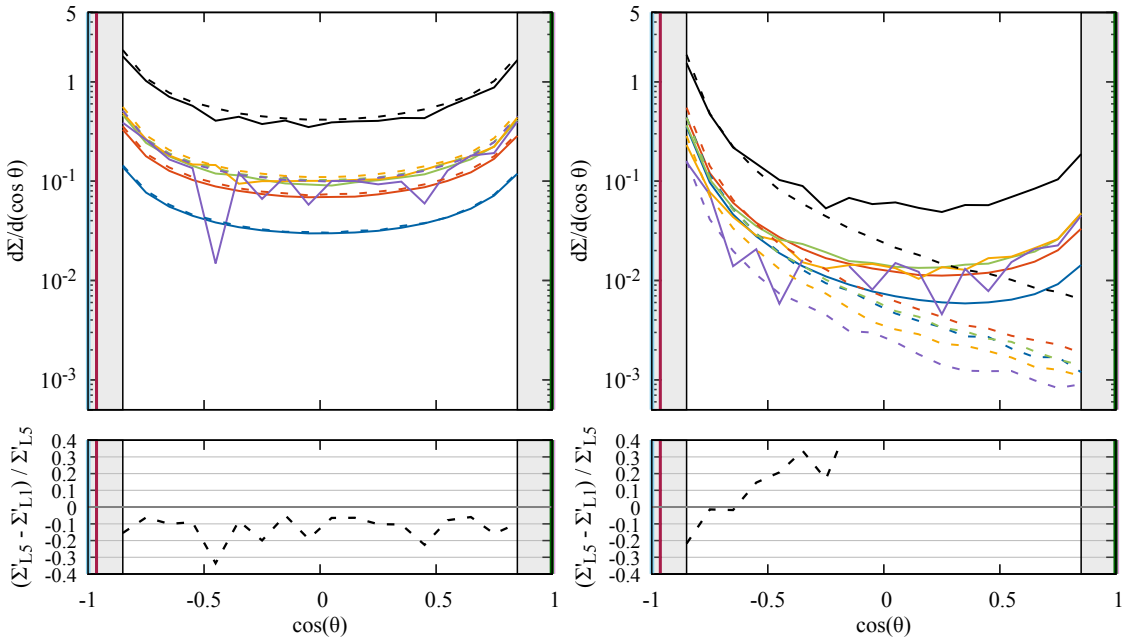


Figure 5.10: The differential cross-sections of the $|10\rangle\langle 01|$ (interference) contribution to the $q\bar{q} \to q\bar{q}$ process, in the fully-symmetric final state configuration, and broken down by multiplicity. The solid curves are the full colour (L5) evolution, and there is no L1 differential cross-section since there are no dipoles to emit from. The locations of the hard jets are marked with vertical lines matching the colours used in Fig. 5.3.



(a) The $|01\rangle\langle 01|$ contribution.

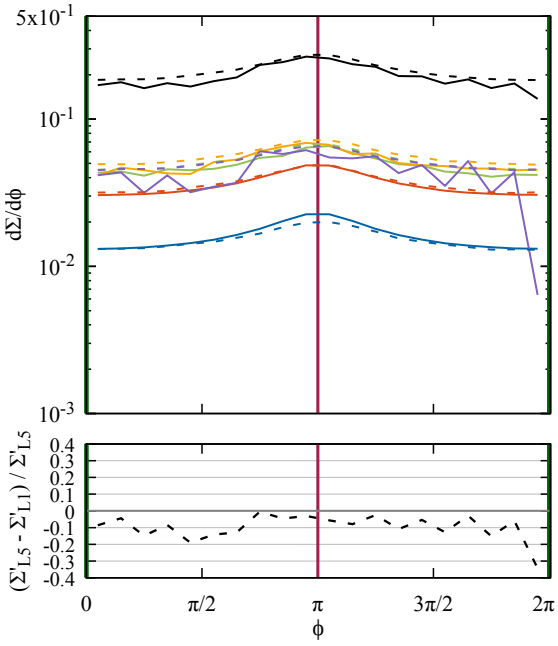
(b) The $|10\rangle\langle 10|$ contribution.



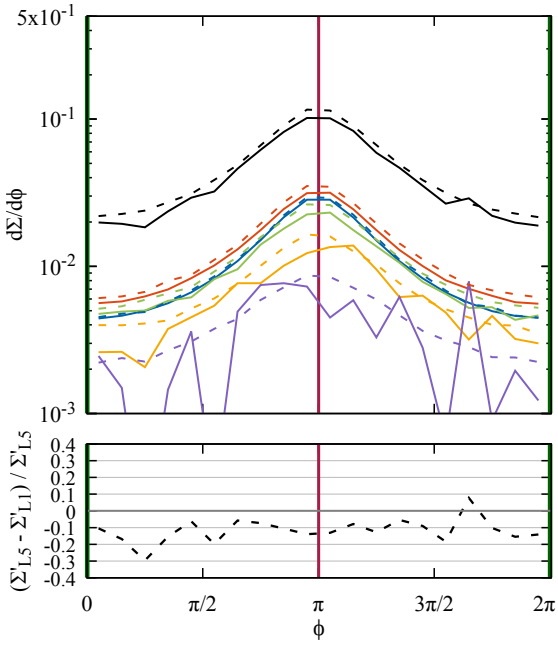
(c) The t-channel gluon exchange contribution.

(d) The s-channel gluon exchange contribution.

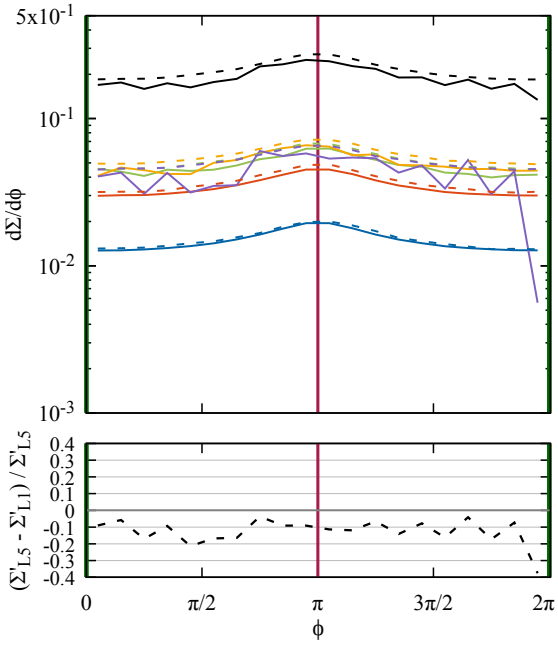
Figure 5.11: The differential cross-section $\mathrm{d}\Sigma/\mathrm{d}(\cos\theta)$ of the different contributions to the $q\bar{q} \to q\bar{q}$ process, in the ϕ-symmetric final state configuration, and broken down by multiplicity. The solid curves are the full colour (L5) evolution, and the dashed curves are the strict LC (L1) evolution. For the L1 curves we start the evolution using the leading-colour approximation to the hard-scatter matrix. The locations of the hard jets are marked with vertical lines matching the colours used in Fig. 5.3. The shaded vertical bars indicate the out-of-gap regions.



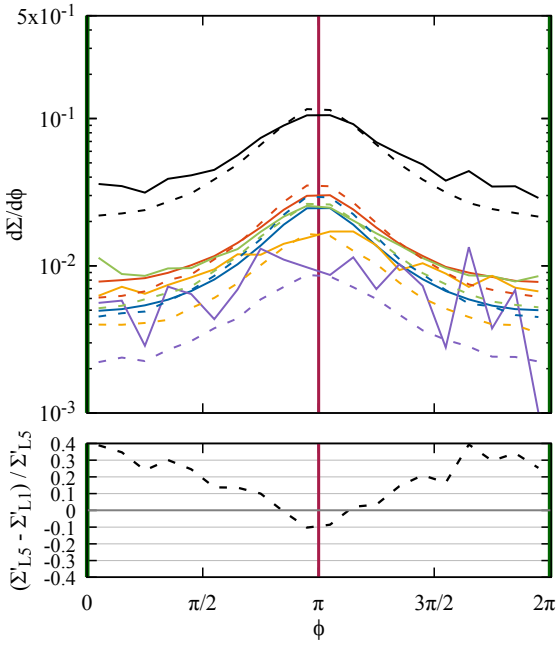
(a) The $|01\rangle\langle 01|$ contribution.



(b) The $|10\rangle\langle 10|$ contribution.

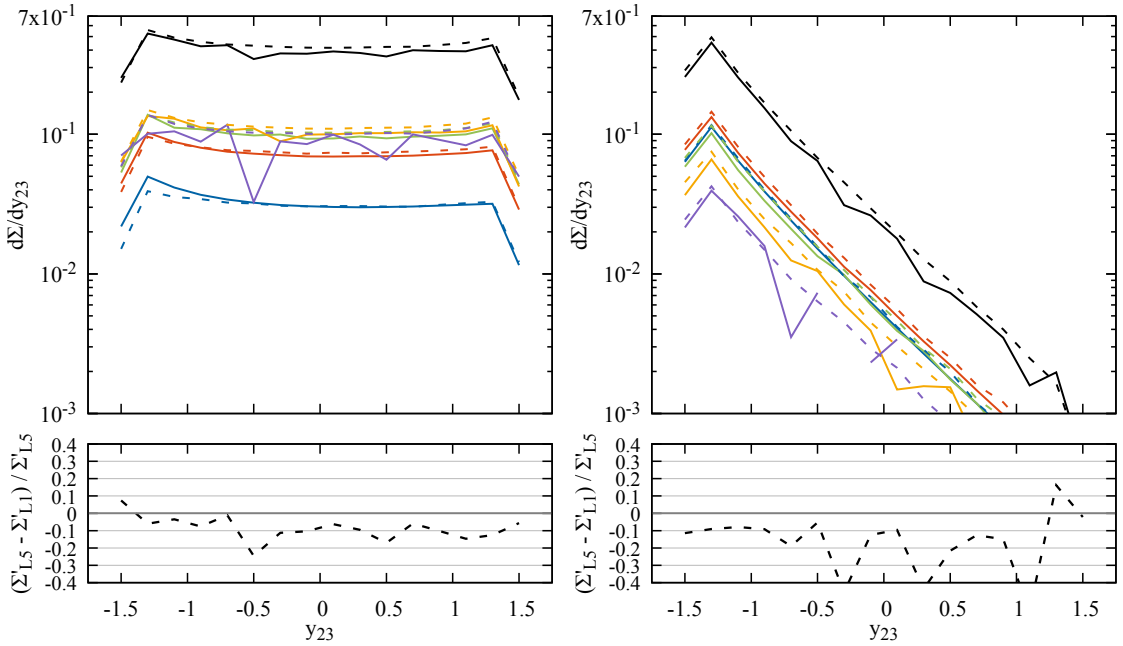


(c) The t-channel gluon exchange contribution.



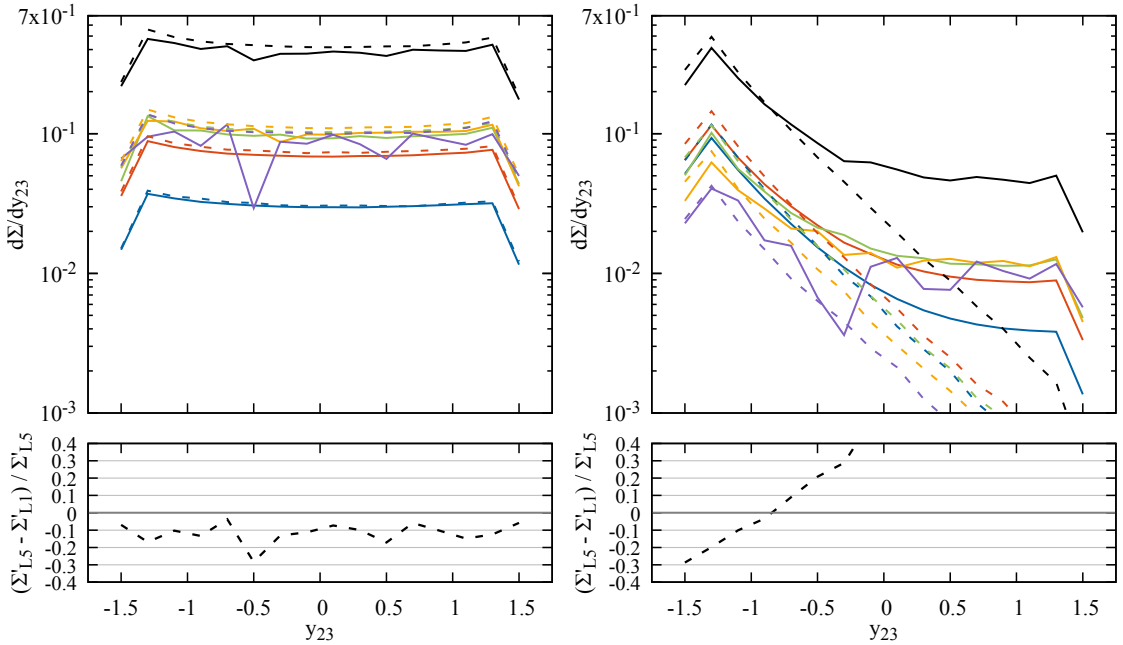
(d) The s-channel gluon exchange contribution.

Figure 5.12: The differential cross-section $\mathrm{d}\Sigma/\mathrm{d}\phi$ of the different contributions to the $q\bar{q} \to q\bar{q}$ process, in the ϕ-symmetric final state configuration, and broken down by multiplicity. The solid curves are the full colour (L5) evolution, and the dashed curves are the strict LC (L1) evolution. For the L1 curves we start the evolution using the leading-colour approximation to the hard-scatter matrix. The locations of the hard jets are marked with vertical lines matching the colours used in Fig. 5.3.



(a) The $|01\rangle\langle 01|$ contribution.

(b) The $|10\rangle\langle 10|$ contribution.



(c) The t-channel gluon exchange contribution.

(d) The s-channel gluon exchange contribution.

Figure 5.13: The differential cross-section $\mathrm{d}\Sigma/\mathrm{d}y_{23}$ of the different contributions to the $q\bar{q} \to q\bar{q}$ process, in the ϕ-symmetric final state configuration, and broken down by multiplicity. The solid curves are the full colour (L5) evolution, and the dashed curves are the strict LC (L1) evolution. For the L1 curves we start the evolution using the leading-colour approximation to the hard-scatter matrix. The locations of the hard jets are marked with vertical lines matching the colours used in Fig. 5.3.

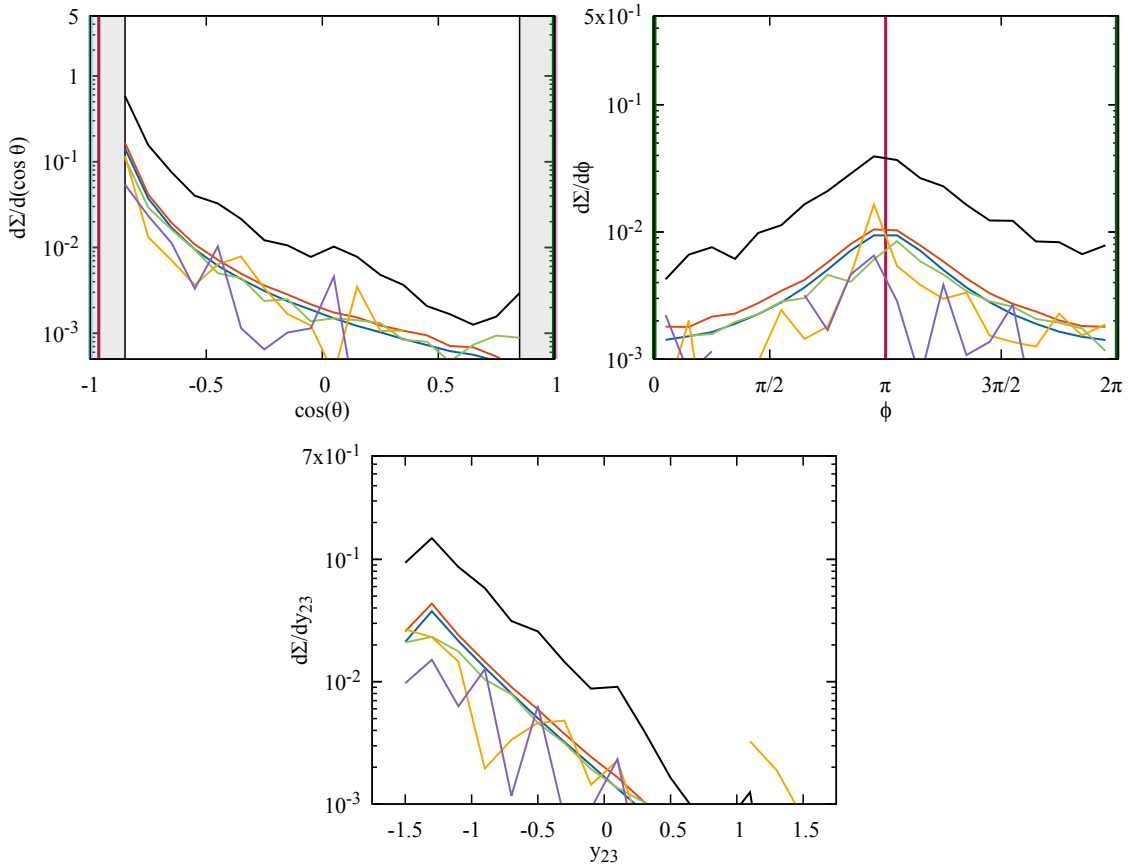
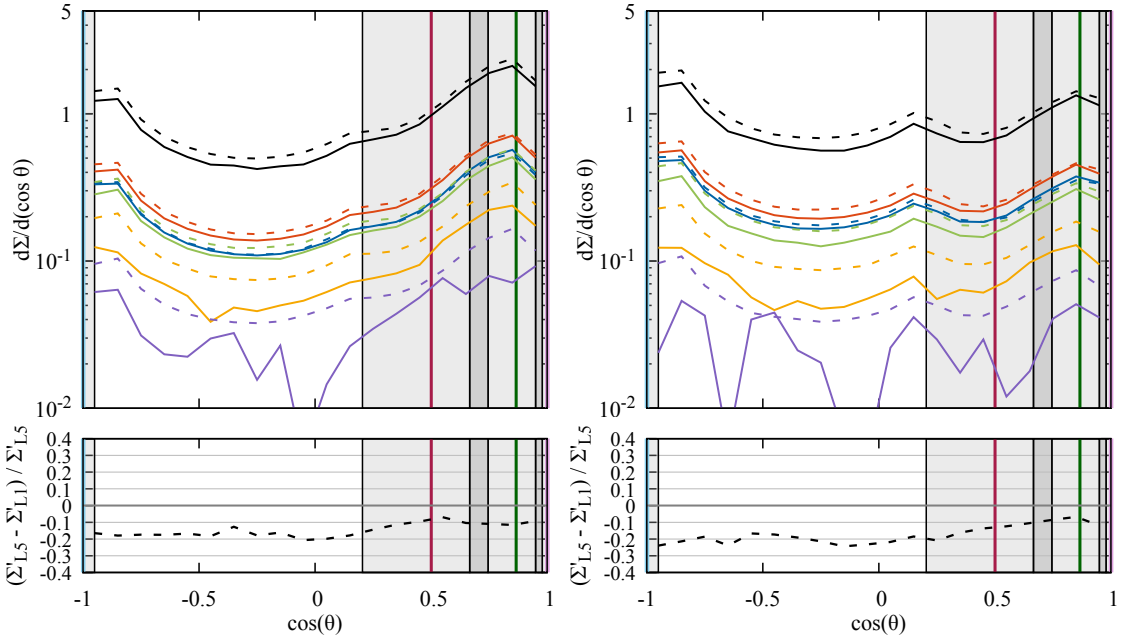
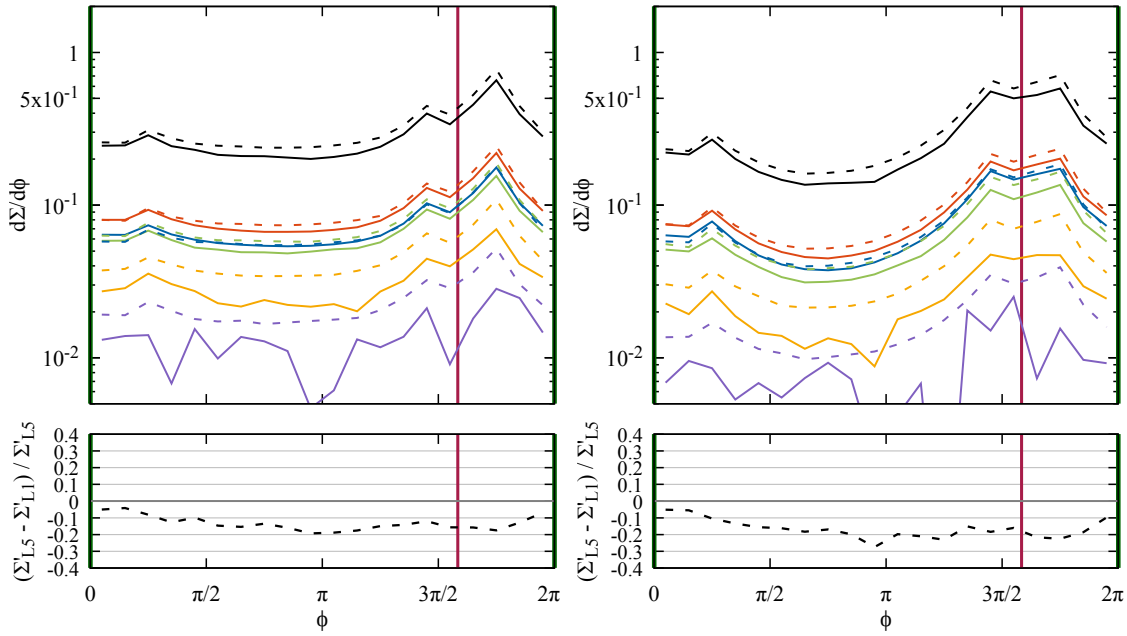


Figure 5.14: The differential cross-sections of the $|10\rangle\langle 01|$ (interference) contribution to the $q\bar{q} \to q\bar{q}$ process, in the ϕ-symmetric final state configuration, and broken down by multiplicity. The solid curves are the full colour (L5) evolution, and there is no L1 cross-section since there are no dipoles to emit from. The locations of the hard jets are marked with vertical lines matching the colours used in Fig. 5.3.



(a) The t-channel gluon exchange contribution. **(b)** The s-channel gluon exchange contribution.

Figure 5.15: The differential cross-section $\mathrm{d}\Sigma/\mathrm{d}(\cos\theta)$ of the different contributions to the $q\bar{q} \to q\bar{q}$ process, in the fully-asymmetric final state configuration, and broken down by multiplicity. The solid curves are the full colour (L5) evolution, and the dashed curves are the strict LC (L1) evolution. For the L1 curves we start the evolution using the leading-colour approximation to the hard-scatter matrix. The locations of the hard jets are marked with vertical lines matching the colours used in Fig. 5.3. The shaded vertical bars indicate the width of the out-of-gap cones around each hard jet. Darker shades indicate an overlap of multiple jet-cones.



(a) The t-channel gluon exchange contribution. **(b)** The s-channel gluon exchange contribution.

Figure 5.16: The differential cross-section $\mathrm{d}\Sigma/\mathrm{d}\phi$ of the different contributions to the $q\bar{q} \to q\bar{q}$ process, in the fully-asymmetric final state configuration, and broken down by multiplicity. The solid curves are the full colour (L5) evolution, and the dashed curves are the strict LC (L1) evolution. For the L1 curves we start the evolution using the leading-colour approximation to the hard-scatter matrix. The locations of the hard jets are marked with vertical lines matching the colours used in Fig. 5.3.

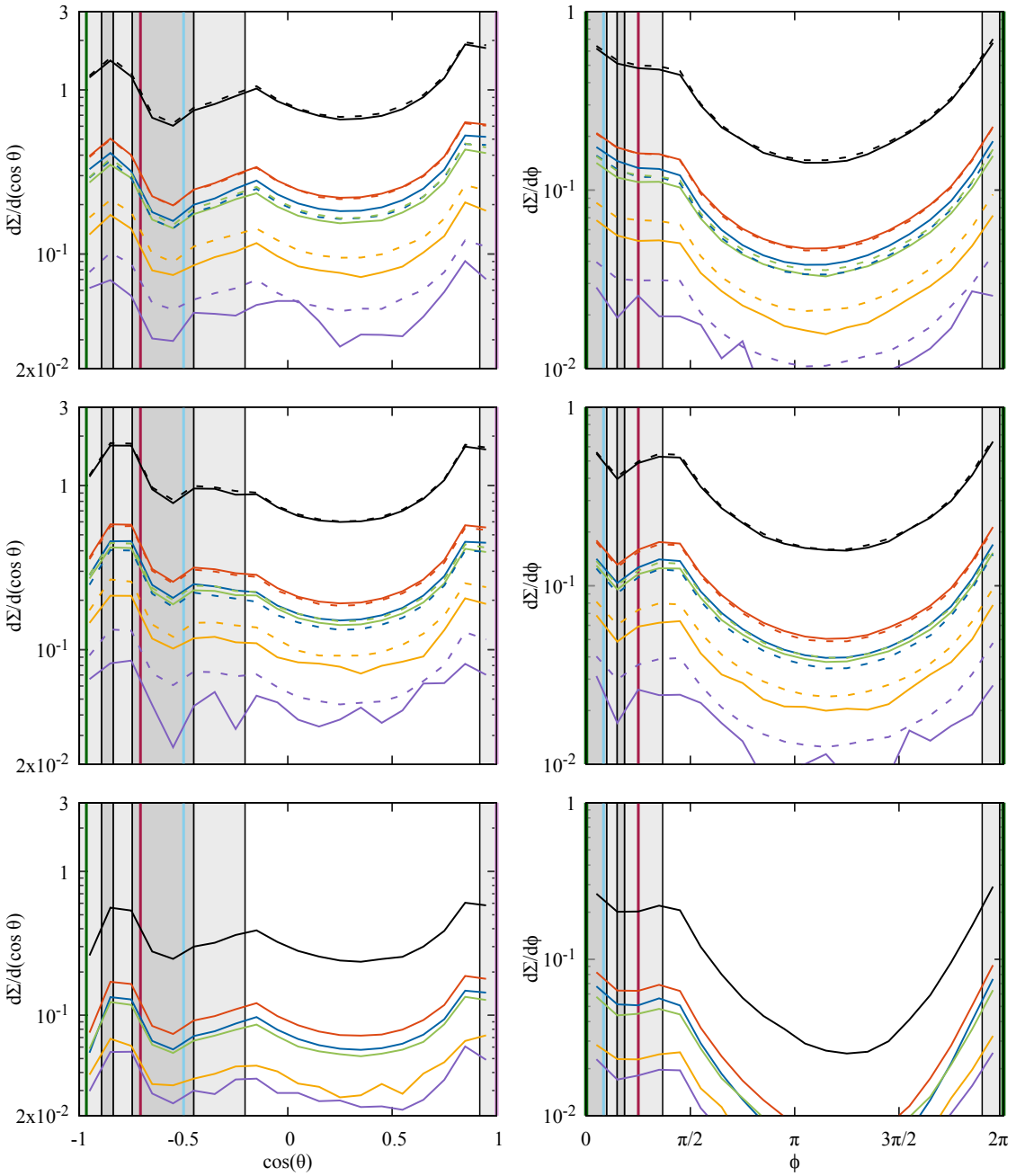


Figure 5.17: The differential cross-sections of the different contributions to the $ZZ \to q\bar{q}q\bar{q}$ process, and broken down by multiplicity. The top row is the $|01\rangle\langle 01|$ contribution, the middle row is the $|10\rangle\langle 10|$ contribution, and the bottom row is the $|10\rangle\langle 01|$ (interference) contribution. The solid curves are the full colour (L5) evolution, and the dashed curves are the leading colour (L1) evolution. For there interference evolution there is no L1 differential cross-section because there are no dipoles to emit from. The locations of the hard jets are marked with vertical lines matching the colours used in Fig. 5.3. The shaded vertical bars indicate the width of the out-of-gap cones around each hard jet. Darker shades indicate an overlap of multiple jet-cones.

Summary and conclusions

Accurate theoretical predictions from QCD are essential for interpreting collider data and uncovering potential signals of new physics. Parton shower algorithms play a critical role in this task by resumming the large logarithms that arise from soft and collinear enhancements at all orders in perturbation theory. However, while traditional parton showers have achieved remarkable success, they cannot incorporate the quantum interference effects which are required for a full-colour treatment of non-global observables.

In recent years, significant progress has been made in computing non-global logarithms beyond the leading-colour limit [110–116]. In this thesis we have continued the development of the amplitude-level parton shower framework, `CVolver`, capable of resumming soft-gluon logarithms at full colour [20, 23, 24]. We have provided a concrete and systematic solution to the long-standing challenge of computing subleading-colour effects within a general-purpose Monte Carlo algorithm.

In Chapter 2, we introduced the general soft-gluon evolution algorithm that underpins our approach. Working in the colour flow basis, we developed a method to track the evolution of scattering amplitudes in colour space, enabling a systematic accounting of colour suppression effects. For virtual corrections, we expanded the Sudakov form factor as an infinite series in $1/N_{\mathrm{c}}$ [18], and for real emissions, we employed the rings and strings formalism to categorise emissions based on their impact on the underlying colour flows [22].

The practical implementation of this algorithm was presented in Chapter 3, where we described the Monte Carlo parton shower `CVolver`. We developed two distinct modes of operation: a dedicated mode tailored to efficiently evaluate wide-angle soft-gluon effects in observables like gaps-between-jets, and a general-purpose event generator mode. We documented the algorithm in detail, including the construction of all relevant sampling distributions and the methods used to guide the evolution toward numerically convergent results. The computational challenge posed by the complexity of the full-colour evolution was considerable and represented the most significant hurdle to overcome in this work.

In Chapter 4, we applied `CVolver` to compute subleading-colour effects in the jet veto cross-section under fixed kinematics. Our findings revealed substantial corrections in many cases, highlighting the importance of including full-colour effects. Interestingly, in a small subset of processes, we observed a non-trivial cancellation of subleading-colour

contributions, in agreement with [110, 112]. These cancellations are non-trivial and point to intricate dynamics in the evolution of colour.

In Chapter 5, we used `CVolver` in event generator mode to study soft radiation patterns differentially across phase-space. These investigations revealed pronounced subleading-colour effects in specific angular regions, underlining the role of colour dynamics in shaping observables sensitive to radiation patterns. Even in cases where the jet veto cross-section at full colour appeared to be well described by leading-colour evolution, we found that differential distributions exposed deviations between full-colour and leading-colour predictions.

Through this work, we have demonstrated that subleading-colour effects in general QCD scattering processes can be computed using amplitude-level parton shower techniques. This opens the door to a range of future investigations. A natural next step is to move beyond the soft-gluon approximation by incorporating hard-collinear emissions and incoming hadrons, since the theoretical framework for their addition to general algorithm has already been developed [21]. Further developments include the implementation of collinear and higher-order anomalous dimensions [19, 104], as well as the study of sophisticated hadronisation models [19]. In addition, the inclusion of collinear physics will allow us to explore super-leading logarithms, which so far have only been studied analytically in specific processes [90, 94–100].

Bibliography

[1] G. Arnison et al. "Experimental observation of isolated large transverse energy electrons with associated missing energy at s=540 GeV". *Physics Letters B* 122.1 (1983), pp. 103–116. DOI: https://doi.org/10.1016/0370-2693(83)91177-2.

[2] P. Bagnaia et al. "Evidence for Z0→e+e− at the CERN pp collider". *Physics Letters B* 129.1 (1983), pp. 130–140. DOI: https://doi.org/10.1016/0370-2693(83)90744-X.

[3] F. Abe et al. "Observation of Top Quark Production in $\bar{p}p$ Collisions with the Collider Detector at Fermilab". *Physical Review Letters* 74.14 (Apr. 1995), pp. 2626–2631. DOI: 10.1103/physrevlett.74.2626.

[4] S. Abachi et al. "Observation of the Top Quark". *Physical Review Letters* 74.14 (Apr. 1995), pp. 2632–2637. DOI: 10.1103/physrevlett.74.2632.

[5] S. Chatrchyan et al. "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC". *Physics Letters B* 716.1 (Sept. 2012), pp. 30–61. DOI: 10.1016/j.physletb.2012.08.021.

[6] G. Aad et al. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". *Physics Letters B* 716.1 (Sept. 2012), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020.

[7] P. W. Higgs. "Broken Symmetries and the Masses of Gauge Bosons". *Phys. Rev. Lett.* 13 (1964). Ed. by J. C. Taylor, pp. 508–509. DOI: 10.1103/PhysRevLett.13.508.

[8] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. "Global Conservation Laws and Massless Particles". *Phys. Rev. Lett.* 13 (1964). Ed. by J. C. Taylor, pp. 585–587. DOI: 10.1103/PhysRevLett.13.585.

[9] "LHC Machine". *JINST* 3 (2008). Ed. by L. Evans and P. Bryant, S08001. DOI: 10.1088/1748-0221/3/08/S08001.

[10] M. Harrison, T. Ludlam, and S. Ozaki. "RHIC project overview". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 499.2 (2003). The Relativistic Heavy Ion Collider Project: RHIC and its Detectors, pp. 235–244. DOI: https://doi.org/10.1016/S0168-9002(02)01937-X.

[11] A. Abada et al. "FCC Physics Opportunities: Future Circular Collider Conceptual Design Report Volume 1". *Eur. Phys. J. C* 79.6 (2019), p. 474. DOI: 10.1140/epjc/s10052-019-6904-3.

[12] J. Bellm et al. "Herwig 7.0/Herwig++ 3.0 release note". *Eur. Phys. J. C* 76.4 (2016), p. 196. DOI: 10.1140/epjc/s10052-016-4018-8. arXiv: 1512.01178 [hep-ph].

[13] G. Bewick et al. *Herwig 7.3 Release Note*. 2024. arXiv: 2312.05175 [hep-ph].

[14] M. Bähr et al. "Herwig++ physics and manual". *The European Physical Journal C* 58.4 (Nov. 2008), pp. 639–707. DOI: 10.1140/epjc/s10052-008-0798-9.

[15] C. Bierlich et al. *A comprehensive guide to the physics and usage of PYTHIA 8.3*. 2022. arXiv: 2203.11601 [hep-ph].

[16] E. Bothmann et al. *Event generation with Sherpa 3*. 2024. arXiv: 2410.22148 [hep-ph].

[17] M. van Beekveld et al. "New Standard for the Logarithmic Accuracy of Parton Showers". *Phys. Rev. Lett.* 134.1 (2025), p. 011901. DOI: 10.1103/PhysRevLett.134.011901. arXiv: 2406.02661 [hep-ph].

[18] S. Plätzer. "Summing Large-N Towers in Colour Flow Evolution". *Eur. Phys. J. C* 74.6 (2014), p. 2907. DOI: 10.1140/epjc/s10052-014-2907-2. arXiv: 1312.2448 [hep-ph].

[19] S. Plätzer. "Colour evolution and infrared physics". *JHEP* 07 (2023), p. 126. DOI: 10.1007/JHEP07(2023)126. arXiv: 2204.06956 [hep-ph].

[20] R. Ángeles Martínez, M. De Angelis, J. R. Forshaw, S. Plätzer, and M. H. Seymour. "Soft gluon evolution and non-global logarithms". *JHEP* 05 (2018), p. 044. DOI: 10.1007/JHEP05(2018)044. arXiv: 1802.08531 [hep-ph].

[21] J. R. Forshaw, J. Holguin, and S. Plätzer. "Parton branching at amplitude level". *JHEP* 08 (2019), p. 145. DOI: 10.1007/JHEP08(2019)145. arXiv: 1905.08686 [hep-ph].

[22] J. R. Forshaw, J. Holguin, and S. Plätzer. "Rings and strings: a basis for understanding subleading colour and QCD coherence beyond the two-jet limit". *JHEP* 05 (2022), p. 190. DOI: 10.1007/JHEP05(2022)190. arXiv: 2112.13124 [hep-ph].

[23] M. De Angelis, J. R. Forshaw, and S. Plätzer. "Resummation and Simulation of Soft Gluon Effects beyond Leading Color". *Phys. Rev. Lett.* 126.11 (2021), p. 112001. DOI: 10.1103/PhysRevLett.126.112001. arXiv: 2007.09648 [hep-ph].

[24] M. De Angelis. "QCD Evolution At Amplitude Level". PhD thesis. Manchester U., 2021.

[25] M. Gell-Mann. "The Eightfold Way: A Theory of strong interaction symmetry" (Mar. 1961). DOI: 10.2172/4008239.

[26] Y. Ne'eman. "Derivation of strong interactions from a gauge invariance". *Nuclear Physics* 26.2 (1961), pp. 222–229. DOI: https://doi.org/10.1016/0029-5582(61)90134-1.

[27] M. Gell-Mann. "A Schematic Model of Baryons and Mesons". *Phys. Lett.* 8 (1964), pp. 214–215. DOI: 10.1016/S0031-9163(64)92001-3.

[28] G. Zweig. "An SU(3) model for strong interaction symmetry and its breaking. Version 2". In: *Developments in the quark theory of hadrons. vol. 1.* Ed. by D. B. Lichtenberg and S. P. Rosen. Feb. 1964, pp. 22–101. DOI: 10.17181/CERN-TH-412.

[29] C. N. Yang and R. L. Mills. "Conservation of Isotopic Spin and Isotopic Gauge Invariance". *Phys. Rev.* 96 (1 Oct. 1954), pp. 191–195. DOI: 10.1103/PhysRev.96.191.

[30] M. Y. Han and Y. Nambu. "Three Triplet Model with Double SU(3) Symmetry". *Phys. Rev.* 139 (1965). Ed. by T. Eguchi, B1006–B1010. DOI: 10.1103/PhysRev.139.B1006.

[31] O. W. Greenberg. "Spin and Unitary Spin Independence in a Paraquark Model of Baryons and Mesons". *Phys. Rev. Lett.* 13 (1964), pp. 598–602. DOI: 10.1103/PhysRevLett.13.598.

[32] S. Weinberg. "Non-Abelian Gauge Theories of the Strong Interactions". *Phys. Rev. Lett.* 31 (7 Aug. 1973), pp. 494–497. DOI: 10.1103/PhysRevLett.31.494.

[33] M. E. Peskin and D. V. Schroeder. *An Introduction to quantum field theory*. Reading, USA: Addison-Wesley, 1995. DOI: 10.1201/9780429503559.

[34] R. K. Ellis, W. J. Stirling, and B. R. Webber. *QCD and collider physics*. Vol. 8. Cambridge University Press, Feb. 2011. DOI: 10.1017/CBO9780511628788.

[35] M. D. Schwartz. *Quantum Field Theory and the Standard Model*. Cambridge University Press, Mar. 2014.

[36] P. A. M. Dirac. "The quantum theory of the electron". *Proc. Roy. Soc. Lond. A* 117 (1928), pp. 610–624. DOI: 10.1098/rspa.1928.0023.

[37] P. A. M. Dirac. "A Theory of Electrons and Protons". *Proc. Roy. Soc. Lond. A* 126.801 (1930), pp. 360–365. DOI: 10.1098/rspa.1930.0013.

[38] D. J. Gross and F. Wilczek. "Ultraviolet Behavior of Nonabelian Gauge Theories". *Phys. Rev. Lett.* 30 (1973). Ed. by J. C. Taylor, pp. 1343–1346. DOI: 10.1103/PhysRevLett.30.1343.

[39] D. J. Gross and F. Wilczek. "Asymptotically Free Gauge Theories. I". *Phys. Rev. D* 8 (10 Nov. 1973), pp. 3633–3652. DOI: 10.1103/PhysRevD.8.3633.

[40] H. D. Politzer. "Reliable Perturbative Results for Strong Interactions?" *Phys. Rev. Lett.* 30 (26 June 1973), pp. 1346–1349. DOI: 10.1103/PhysRevLett.30.1346.

[41] H. Georgi. *Lie algebras in particle physics*. en. London, England: CRC Press, May 2019.

[42] M. Gell-Mann. "Symmetries of Baryons and Mesons". *Phys. Rev.* 125 (3 Feb. 1962), pp. 1067–1084. DOI: 10.1103/PhysRev.125.1067.

[43] G. 't Hooft. "Renormalization of Massless Yang-Mills Fields". *Nucl. Phys. B* 33 (1971), pp. 173–199. DOI: 10.1016/0550-3213(71)90395-6.

[44] L. Faddeev and V. Popov. "Feynman diagrams for the Yang-Mills field". *Physics Letters B* 25.1 (1967), pp. 29–30. DOI: https://doi.org/10.1016/0370-2693(67)90067-6.

[45] G. 't Hooft. "An algorithm for the poles at dimension four in the dimensional regularization procedure". *Nuclear Physics B* 62 (1973), pp. 444–460. DOI: https://doi.org/10.1016/0550-3213(73)90263-0.

[46] G. 't Hooft. "Renormalizable Lagrangians for massive Yang-Mills fields". *Nuclear Physics B* 35.1 (1971), pp. 167–188. DOI: https://doi.org/10.1016/0550-3213(71)90139-8.

[47] P. A. M. Dirac. "The quantum theory of the emission and absorption of radiation". en. *Proc. R. Soc. Lond. A Math. Phys. Sci.* 114.767 (Mar. 1927), pp. 243–265.

[48] M. Srednicki. *Quantum Field Theory*. Cambridge, England: Cambridge University Press, Jan. 2007.

[49] A. Zee. *Quantum field theory in a nutshell*. en. 2nd ed. In a Nutshell. Princeton, NJ: Princeton University Press, Feb. 2010.

[50] I. V. Tyutin. *Gauge Invariance in Field Theory and Statistical Physics in Operator Formalism*. 2008. arXiv: 0812.0580 [hep-th].

[51] C. Becchi, A. Rouet, and R. Stora. "Renormalization of gauge theories". *Annals of Physics* 98.2 (1976), pp. 287–321. DOI: https://doi.org/10.1016/0003-4916(76)90156-1.

[52] M. Veltman and G. 't Hooft. "Combinatorics of gauge fields". *Nuclear Physics B* 50.1 (1972), pp. 318–353. DOI: https://doi.org/10.1016/S0550-3213(72)80021-X.

[53] D. Pritchard and W. Stirling. "QCD calculations in the light-cone gauge". *Nuclear Physics B* 165.2 (1980), pp. 237–268. DOI: https://doi.org/10.1016/0550-3213(80)90086-3.

[54] S. Mandelstam. "Determination of the Pion-Nucleon Scattering Amplitude from Dispersion Relations and Unitarity. General Theory". *Phys. Rev.* 112 (4 Nov. 1958), pp. 1344–1360. DOI: 10.1103/PhysRev.112.1344.

[55] S. Weinberg. "A Model of Leptons". *Phys. Rev. Lett.* 19 (1967), pp. 1264–1266. DOI: 10.1103/PhysRevLett.19.1264.

[56] F. Maltoni, K. Paul, T. Stelzer, and S. Willenbrock. "Color-flow decomposition of QCD amplitudes". *Physical Review D* 67.1 (Jan. 2003). DOI: 10.1103/physrevd.67.014026.

[57] W. Kilian, T. Ohl, J. Reuter, and C. Speckner. "QCD in the color-flow representation". *Journal of High Energy Physics* 2012.10 (Oct. 2012). DOI: 10.1007/jhep10(2012)022.

[58] S. Catani and M. Seymour. "A general algorithm for calculating jet cross sections in NLO QCD". *Nuclear Physics B* 485.1–2 (Feb. 1997), pp. 291–419. DOI: 10.1016/s0550-3213(96)00589-5.

[59] M. Lévy and J. Sucher. "Eikonal Approximation in Quantum Field Theory". *Phys. Rev.* 186 (5 Oct. 1969), pp. 1656–1670. DOI: 10.1103/PhysRev.186.1656.

[60] J. Holguin. "Developing an amplitude level approach to radiative processes in quantum chromodynamics". PhD thesis. University of Manchester, Manchester U., 2021.

[61] A. Bassetto, M. Ciafaloni, and G. Marchesini. "Jet structure and infrared sensitive quantities in perturbative QCD". *Physics Reports* 100.4 (1983), pp. 201–272. DOI: https://doi.org/10.1016/0370-1573(83)90083-2.

[62] S. Catani and M. Grazzini. "The soft-gluon current at one-loop order". *Nuclear Physics B* 591.1–2 (Dec. 2000), pp. 435–454. DOI: 10.1016/s0550-3213(00)00572-1.

[63] Z. Bern, V. Del Duca, W. B. Kilgore, and C. R. Schmidt. "Infrared behavior of one-loop QCD amplitudes at next-to-next-to-leading order". *Physical Review D* 60.11 (Oct. 1999). DOI: 10.1103/physrevd.60.116001.

[64] G. Altarelli and G. Parisi. "Asymptotic freedom in parton language". *Nuclear Physics B* 126.2 (1977), pp. 298–318. DOI: https://doi.org/10.1016/0550-3213(77)90384-4.

[65] Y. L. Dokshitzer, V. A. Khoze, A. H. Mueller, and S. I. Troian. *Basics of perturbative QCD*. 1991.

[66] T. Kinoshita. "Mass singularities of Feynman amplitudes". *J. Math. Phys.* 3 (1962), pp. 650–677. DOI: 10.1063/1.1724268.

[67] T. D. Lee and M. Nauenberg. "Degenerate Systems and Mass Singularities". *Phys. Rev.* 133 (1964). Ed. by G. Feinberg, B1549–B1562. DOI: 10.1103/PhysRev.133.B1549.

[68] F. Bloch and A. Nordsieck. "Note on the Radiation Field of the Electron". *Phys. Rev.* 52 (2 July 1937), pp. 54–59. DOI: 10.1103/PhysRev.52.54.

[69] R. D. Field. *Applications of Perturbative QCD*. Vol. 77. 1989.

[70] G. 't Hooft and M. J. G. Veltman. "Regularization and Renormalization of Gauge Fields". *Nucl. Phys. B* 44 (1972), pp. 189–213. DOI: 10.1016/0550-3213(72)90279-9.

[71] G. 't Hooft. "Dimensional regularization and the renormalization group". *Nuclear Physics B* 61 (1973), pp. 455–468. DOI: https://doi.org/10.1016/0550-3213(73)90376-3.

[72] J. C. Collins. *Renormalization : An Introduction to Renormalization, the Renormalization Group and the Operator-Product Expansion*. Vol. 26. Cambridge Monographs on Mathematical Physics. Cambridge: Cambridge University Press, 1984. DOI: 10.1017/9781009401807.

[73] M. Aaboud et al. "Search for the Standard Model Higgs boson produced by vector-boson fusion and decaying to bottom quarks in $\sqrt{s} = 8$ TeV pp collisions with the ATLAS detector". *Journal of High Energy Physics* 2016.11 (Nov. 2016). DOI: 10.1007/jhep11(2016)112.

[74] J. C. Collins, D. E. Soper, and G. Sterman. "Soft gluons and factorization". *Nuclear Physics B* 308.4 (1988), pp. 833–856. DOI: https://doi.org/10.1016/0550-3213(88)90130-7.

[75] G. Oderda and G. Sterman. "Energy and Color Flow in Dijet Rapidity Gaps". *Physical Review Letters* 81.17 (Oct. 1998), pp. 3591–3594. DOI: 10.1103/physrevlett.81.3591.

[76] G. Oderda. "Dijet rapidity gaps in photoproduction from perturbative QCD". *Physical Review D* 61.1 (Nov. 1999). DOI: 10.1103/physrevd.61.014004.

[77] Y. L. Dokshitzer and G. Marchesini. "On large angle multiple gluon radiation". *Journal of High Energy Physics* 2003.03 (Mar. 2003), pp. 040–040. DOI: 10.1088/1126-6708/2003/03/040.

[78] J. Botts and G. Sterman. "Hard elastic scattering in QCD: Leading behavior". *Nuclear Physics B* 325.1 (1989), pp. 62–100. DOI: https://doi.org/10.1016/0550-3213(89)90372-6.

[79] N. Kidonakis, G. Oderda, and G. Sterman. "Evolution of color exchange in QCD hard scattering". *Nuclear Physics B* 531.1 (1998), pp. 365–402. DOI: https://doi.org/10.1016/S0550-3213(98)00441-6.

[80] R. B. Appleby and M. H. Seymour. "Non-global logarithms in inter-jet energy flow with kt clustering requirement". *Journal of High Energy Physics* 2002.12 (Dec. 2002), pp. 063–063. DOI: 10.1088/1126-6708/2002/12/063.

[81] R. B. Appleby and M. H. Seymour. "The resummation of inter-jet energy flow for gaps-between-jets processes at HERA". *Journal of High Energy Physics* 2003.09 (Sept. 2003), pp. 056–056. DOI: 10.1088/1126-6708/2003/09/056.

[82] J. R. Forshaw, A. Kyrieleis, and M. H. Seymour. "Gaps between Jets in the high energy limit". *Journal of High Energy Physics* 2005.06 (June 2005), pp. 034–034. DOI: 10.1088/1126-6708/2005/06/034.

[83] A. Banfi, G. Marchesini, and G. Smye. "Away-from-jet energy flow". *Journal of High Energy Physics* 2002.08 (Aug. 2002), pp. 006–006. DOI: 10.1088/1126-6708/2002/08/006.

[84] A. Banfi, G. P. Salam, and G. Zanderighi. "Principles of general final-state resummation and automated implementation". *Journal of High Energy Physics* 2005.03 (Mar. 2005), pp. 073–073. DOI: 10.1088/1126-6708/2005/03/073.

[85] Y. Dokshitzer and G. Marchesini. "Soft gluons at large angles in hadron collisions". *Journal of High Energy Physics* 2006.01 (Jan. 2006), pp. 007–007. DOI: 10.1088/1126-6708/2006/01/007.

[86] M. Dasgupta and G. Salam. "Resummation of non-global QCD observables". *Physics Letters B* 512.3–4 (July 2001), pp. 323–330. DOI: 10.1016/s0370-2693(01)00725-0.

[87] J. R. Forshaw, A. Kyrieleis, and M. H. Seymour. "Super-leading logarithms in non-global observables in QCD". *JHEP* 08 (2006), p. 059. DOI: 10.1088/1126-6708/2006/08/059. arXiv: hep-ph/0604094.

[88] 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), p. 128. DOI: 10.1088/1126-6708/2008/09/128. arXiv: 0808.1269 [hep-ph].

[89] J. R. Forshaw, M. H. Seymour, and A. Siódmok. "On the breaking of collinear factorization in QCD". *Journal of High Energy Physics* 2012.11 (Nov. 2012). DOI: 10.1007/jhep11(2012)066.

[90] J. Keates and M. H. Seymour. "Super-leading logarithms in non-global observables in QCD: fixed order calculation". *Journal of High Energy Physics* 2009.04 (Apr. 2009), p. 040. DOI: 10.1088/1126-6708/2009/04/040.

[91] M. D. Schwartz, K. Yan, and H. X. Zhu. "Collinear factorization violation and effective field theory". *Phys. Rev. D* 96 (5 Sept. 2017), p. 056005. DOI: 10.1103/PhysRevD.96.056005.

[92] J. R. Forshaw and J. Holguin. "Coulomb gluons will generally destroy coherence". en. *J. High Energy Phys.* 2021.12 (Dec. 2021).

[93] J. R. Forshaw, M. H. Seymour, and A. Siódmok. "On the breaking of collinear factorization in QCD". *Journal of High Energy Physics* 2012.11 (Nov. 2012). DOI: 10.1007/jhep11(2012)066.

[94] T. Becher, P. Hager, G. Martinelli, M. Neubert, D. Schwienbacher, and M. Stillger. "Super-leading logarithms in pp $\to$ 2 jets". *JHEP* 01 (2025), p. 171. DOI: 10.1007/JHEP01(2025)171. arXiv: 2411.12742 [hep-ph].

[95] T. Becher, M. Neubert, D. Y. Shao, and M. Stillger. "Factorization of non-global LHC observables and resummation of super-leading logarithms". *Journal of High Energy Physics* 2023.12 (Dec. 2023). DOI: 10.1007/jhep12(2023)116.

[96] P. Böer, P. Hager, M. Neubert, M. Stillger, and X. Xu. *Renormalization-Group Improved Resummation of Super-Leading Logarithms*. 2024. arXiv: 2405.05305 [hep-ph].

[97] T. Becher, M. Neubert, and D. Y. Shao. "Resummation of Super-Leading Logarithms". *Physical Review Letters* 127.21 (Nov. 2021). DOI: 10.1103/physrevlett.127.212002.

[98] T. Becher, P. Hager, S. Jaskiewicz, M. Neubert, and D. Schwienbacher. *Factorization restoration through Glauber gluons*. 2024. arXiv: 2408.10308 [hep-ph].

[99] P. Böer, M. Neubert, and M. Stillger. "Glauber phases in non-global LHC observables: resummation for quark-initiated processes". *Journal of High Energy Physics* 2023.10 (Oct. 2023). DOI: 10.1007/jhep10(2023)075.

[100] P. Böer, P. Hager, M. Neubert, M. Stillger, and X. Xu. *Resummation of Glauber Phases in Non-Global LHC Observables for Large N_c*. 2024. arXiv: 2407.01691 [hep-ph].

[101] J. Alcock-Zeilinger and H. Weigert. "A simple counting argument of the irreducible representations of SU(N) on mixed product spaces". *Journal of Algebraic Combinatorics* 50.3 (2018), pp. 281–291. DOI: 10.1007/s10801-018-0853-z.

[102] Y. Dokshitzer and G. Marchesini. "Hadron collisions and the fifth form factor". *Physics Letters B* 631.3 (Dec. 2005), pp. 118–125. DOI: 10.1016/j.physletb.2005.10.009.

[103] J. R. Forshaw, S. Plätzer, and F. Torre González. *Exact colour evolution for jet observables*. 2025. arXiv: 2502.12133 [hep-ph].

[104] S. Plätzer and I. Ruffa. "Towards colour flow evolution at two loops". *Journal of High Energy Physics* 2021.6 (June 2021). DOI: 10.1007/jhep06(2021)007.

[105] M. De Angelis, J. R. Forshaw, P. Kirchgaesser, S. Plätzer, and F. Torre González. "`CVolver` soft gluon plugin". In preparation. 2025.

[106] S. Plätzer and M. Sjödahl. "The Sudakov Veto Algorithm Reloaded". *Eur. Phys. J. Plus* 127 (2012), p. 26. DOI: 10.1140/epjp/i2012-12026-x. arXiv: 1108.6180 [hep-ph].

[107] The HDF Group. *Hierarchical Data Format, version 5*.

[108] B. Fortner. "HDF: The hierarchical data format". *Dr Dobb's J Software Tools Prof Program* 23.5 (1998), p. 42.

[109] S. Plätzer. "myStatistics Histogramming Package". In preparation. 2025.

[110] Y. Hatta and T. Ueda. "Resummation of non-global logarithms at finite N_c". *Nucl. Phys. B* 874 (2013), pp. 808–820. DOI: 10.1016/j.nuclphysb.2013.06.021. arXiv: 1304.6930 [hep-ph].

[111] K. Hamilton, R. Medves, G. P. Salam, L. Scyboz, and G. Soyez. "Colour and logarithmic accuracy in final-state parton showers". *Journal of High Energy Physics* 2021.3 (Mar. 2021). DOI: 10.1007/jhep03(2021)041.

[112] Y. Hatta and T. Ueda. "Non-global logarithms in hadron collisions at $N_c = 3$". *Nucl. Phys. B* 962 (2021), p. 115273. DOI: 10.1016/j.nuclphysb.2020.115273. arXiv: 2011.04154 [hep-ph].

[113] Z. Nagy and D. E. Soper. "Parton shower evolution with subleading color". *JHEP* 06 (2012), p. 044. DOI: 10.1007/JHEP06(2012)044. arXiv: 1202.4496 [hep-ph].

[114] Z. Nagy and D. E. Soper. "Effects of subleading color in a parton shower". *JHEP* 07 (2015), p. 119. DOI: 10.1007/JHEP07(2015)119. arXiv: 1501.00778 [hep-ph].

[115] Z. Nagy and D. E. Soper. "Parton showers with more exact color evolution". *Phys. Rev. D* 99.5 (2019), p. 054009. DOI: 10.1103/PhysRevD.99.054009. arXiv: 1902.02105 [hep-ph].

[116] J. Holguin, J. R. Forshaw, and S. Plätzer. "Improvements on dipole shower colour". *The European Physical Journal C* 81.4 (Apr. 2021). DOI: 10.1140/epjc/s10052-021-09145-1.

[117] G. Peter Lepage. "A new algorithm for adaptive multidimensional integration". *Journal of Computational Physics* 27.2 (1978), pp. 192–203. DOI: https://doi.org/10.1016/0021-9991(78)90004-9.

[118] C. Sanderson and R. Curtin. "Armadillo: A template-based C++ library for linear algebra". *Journal of Open Source Software* 1 (July 2016), p. 26. DOI: 10.21105/joss.00026.

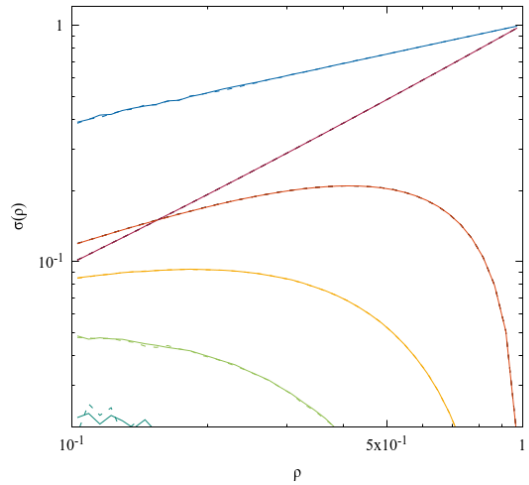
[119] S. Plätzer. "matrices2 Library". In preparation. 2025.

Appendices

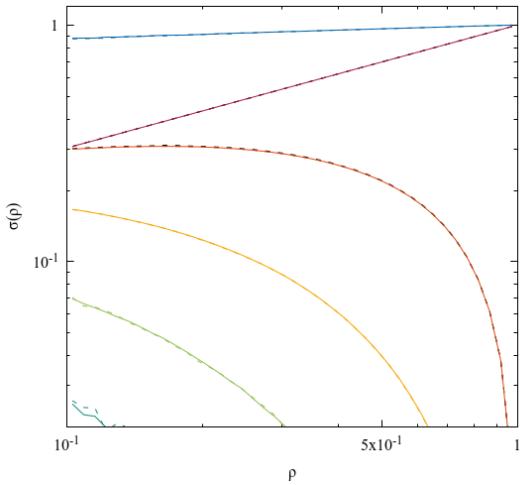
Appendix A

Validation plots

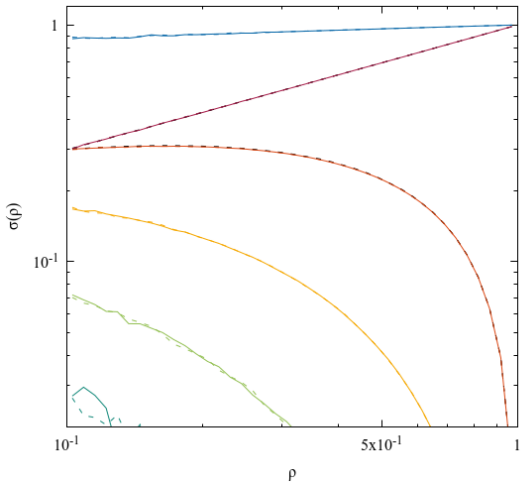
We compile a small selection of validation plots, with the purpose of verifying the soft gluon algorithm and the `CVolver` implementation.



(a) The $|01\rangle\langle 01|$ contribution

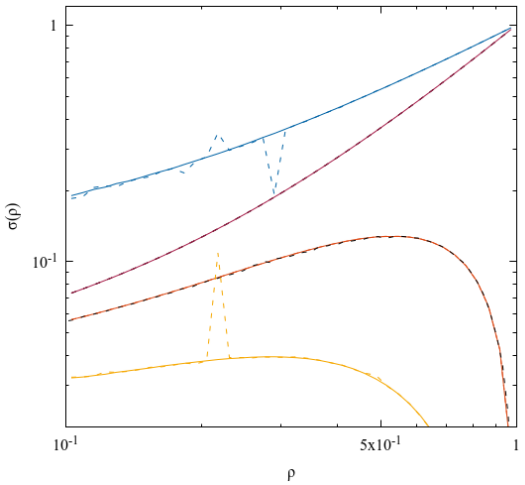


(b) The $|10\rangle\langle 10|$ contribution

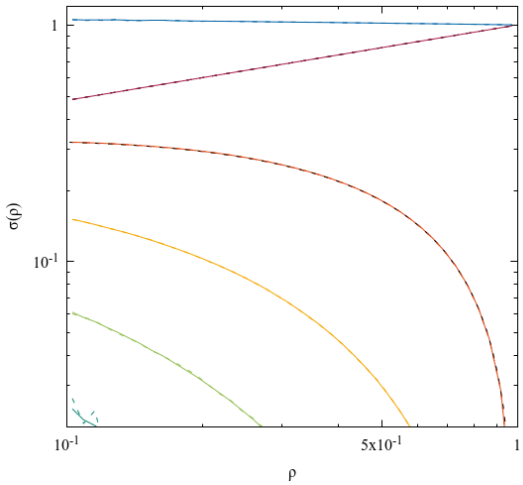


(c) The $|10\rangle\langle 01|$ contribution

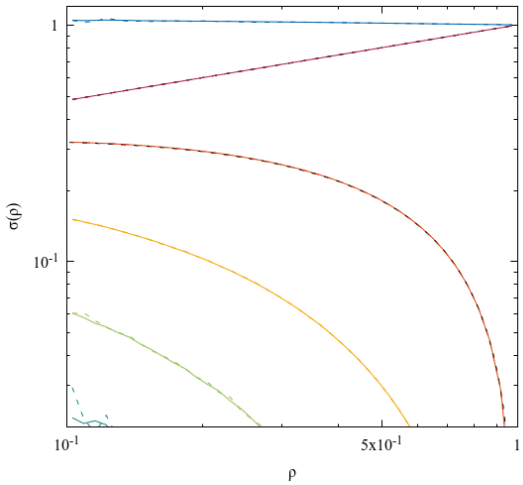
Figure A.1: Gap veto cross-section, broken down by multiplicity, for $q\bar{q} \to q\bar{q}$ in the symmetric kinematic configuration defined in Chapter 4. The solid curves correspond to $d = 2$, and the dashed curves correspond to $d = 3$. This is evidence that $d = 2$ is enough for the Sudakov expansion to converge for this process. Two additional black-dashed curves are overlaid on the 0 and 1 emission curves, respectively. These are independent semi-analytic calculations, performed with the `GSL` Vegas integrator [117], the `Armadillo` package for exponentiating the anomalous dimension [118], and the `Matrices2` library for evaluating matrix elements in the colour flow basis [119]. They show perfect agreement with `CVolver`, which validates the implementation of the evolution and also the convergence of the Sudakov expansion. These semi-analytic cross-checks were performed for every process presented in this thesis, up to 1 emission. The only exceptions are $gg \to gg$ and some kinematic configurations of $qg \to qg$, for which the matrix exponentiation does not converge.



(a) The $|01\rangle\langle 01|$ contribution

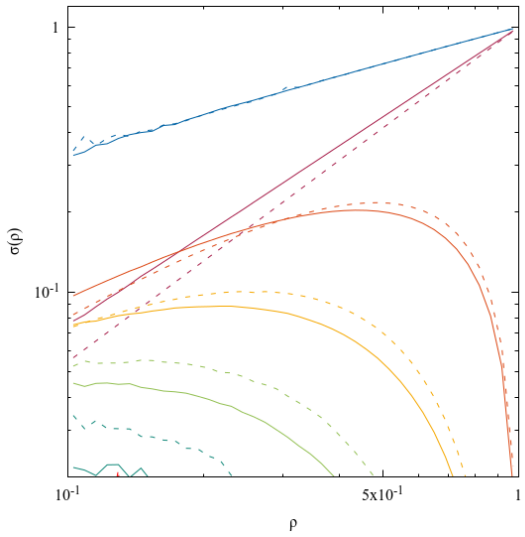


(b) The $|10\rangle\langle 10|$ contribution

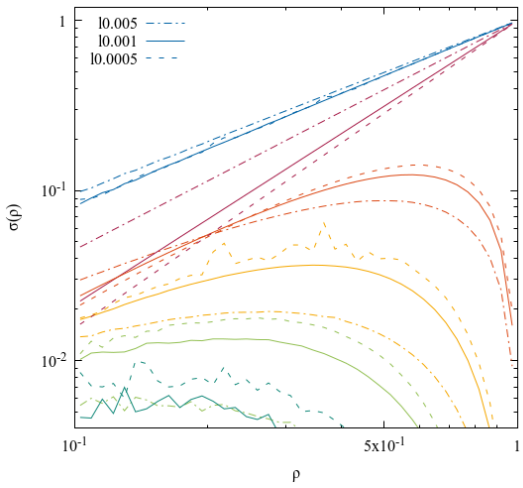


(c) The $|10\rangle\langle 01|$ contribution

Figure A.2: Gap veto cross-section, broken down by multiplicity, for $q\bar{q} \to q\bar{q}$ in the asymmetric kinematic configuration defined in Chapter 4. The solid curves correspond to $d = 2$, and the dashed curves correspond to $d = 3$. This is evidence that $d = 2$ is enough for the Sudakov expansion to converge for this process. The semi-analytic black dashed curves all agree with `CVolver`.

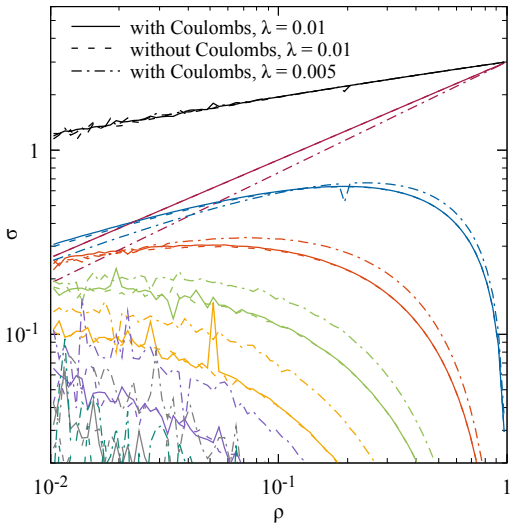


(a) Symmetric kinematic configuration. The solid curves are $\lambda = 0.01$, and the dashed are $\lambda = 0.005$.

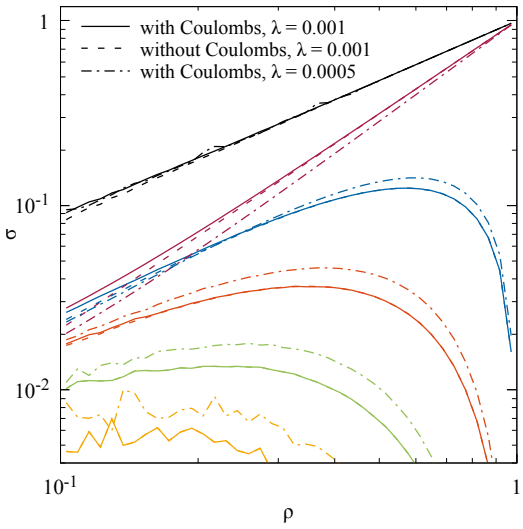


(b) Asymmetric kinematic configuration. The solid curves are $\lambda = 0.001$, the dashed are $\lambda = 0.0005$, and the dash-dotted are $\lambda = 0.005$.

Figure A.3: Gap veto cross-section, broken down by multiplicity, for t-channel gluon exchange $q\bar{q} \to q\bar{q}$ in the kinematic configurations defined in Chapter 4. Testing independence of the collinear cutoff λ, defined in Eq. (3.27). The collinear cutoff required for independence depends on the kinematics of the process: $\lambda = 0.005$ is too big in the asymmetric kinematics, but is small enough for the symmetric case.

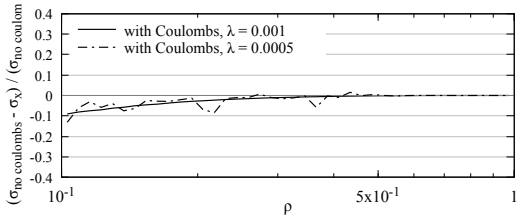


(a) $Z \to q\bar{q}$

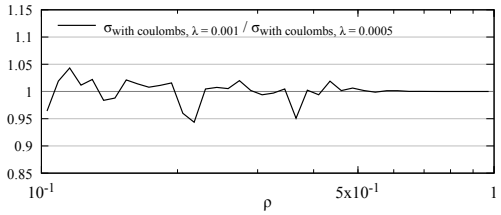


(b) t-channel gluon exchange $q\bar{q} \to q\bar{q}$, asymmetric configuration

Figure A.4: Testing the impact of including Coulomb gluons. In $Z \to q\bar{q}$, which has a colourless initial state, the imaginary parts cancel. The cross-section remains collinear cutoff independent with Coulombs included. In $q\bar{q} \to q\bar{q}$ including the Coulombs has a noticeable effect. The existence of super-leading logarithms means we expect the cross-section to become collinear cutoff dependent: the effect is small for $\rho > 0.1$. Semi-analytic tests were also performed with Coulomb gluons, and found to agree with `CVolver`.



(a) Including Coulombs induces a 10% effect at $\rho = 0.1$



(b) It is necessary to go lower in ρ to confirm the collinear cutoff sensitivity.

Figure A.5: Effect of Coulomb gluons in t-channel gluon exchange $q\bar{q} \to q\bar{q}$ in the asymmetric configuration.

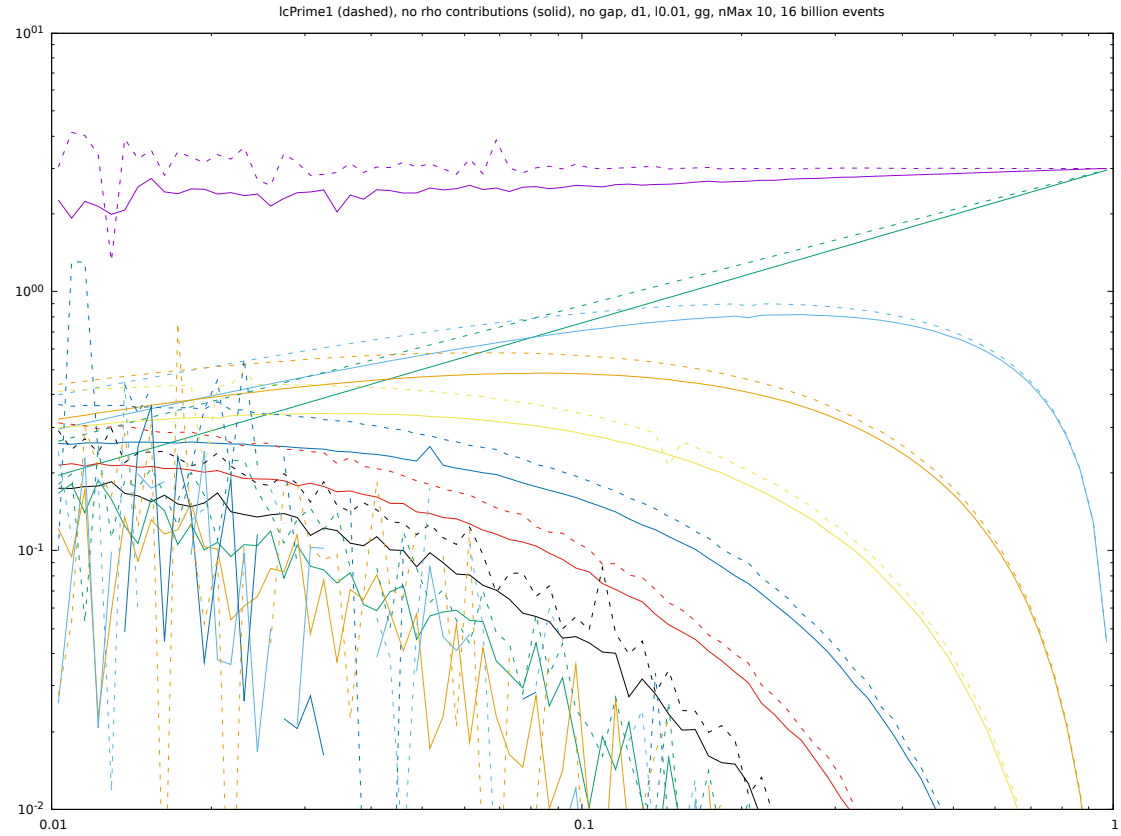


Figure A.6: Unitarity test, the gap is removed and all emissions are accepted in dedicated mode. The process is $Z \to q\bar{q}$, broken down by multiplicity. The dashed lines are full colour, and the full colour total cross-section remains constant. The solid lines are full colour except for the ρ contributions in the virtuals, which are removed. It fails the unitarity test.