Karlsruher Institut für Technologie

Non-perturbative Models for the Simulation of Hadronic Collisions at the LHC

Zur Erlangung des akademischen Grades eines

DOKTORS DER NATURWISSENSCHAFTEN

von der KIT-Fakultät für Physik
des Karlsruher Instituts für Technologie (KIT)

genehmigte

DISSERTATION

von

M.Sc. Patrick Kirchgaeßer

aus Ludwigsburg

Referent: PD Dr. S. Gieseke (KIT, Karlsruhe)

Korreferent: Prof. Dr. Dieter Zeppenfeld (KIT, Karlsruhe)

Tag der mündlichen Prüfung: 24. April 2020

Eidesstattliche Versicherung gemäß § 13 Absatz 2 Ziffer 3 der Promotionsordnung des Karlsruher Instituts für Technologie (KIT) für die KIT-Fakultät für Physik:

1. Bei der eingereichten Dissertation zu dem Thema

   Non-perturbative Models for the Simulation of Hadronic Collisions at the LHC

   handelt es sich um meine eigenständig erbrachte Leistung.

2. Ich habe nur die angegebenen Quellen und Hilfsmittel benutzt und mich keiner unzulässigen Hilfe Dritter bedient. Insbesondere habe ich wörtlich oder sinngemäß aus anderen Werken übernommene Inhalte als solche kenntlich gemacht.

3. Die Arbeit oder Teile davon habe ich bislang nicht an einer Hochschule des In- oder Auslands als Bestandteil einer Prüfungs- oder Qualifikationsleistung vorgelegt.

4. Die Richtigkeit der vorstehenden Erklärungen bestätige ich.

5. Die Bedeutung der eidesstattlichen Versicherung und die strafrechtlichen Folgen einer unrichtigen oder unvollständigen eidesstattlichen Versicherung sind mir bekannt.

Ich versichere an Eides statt, dass ich nach bestem Wissen die reine Wahrheit erklärt und nichts verschwiegen habe.

Karlsruhe, den 08. April 2020

. . . . . .

(Patrick Kirchgaeßer)

"Engage!"

Cpt. Jean Luc Picard
(Star Trek - The Next Generation)

Abstract

Non-perturbative models are an essential part in the simulation of high-energy hadronic collisions. In this thesis we study, implement, and improve several non-perturbative models in the Monte Carlo event generator Herwig 7. We first develop a new model for colour reconnection which allows the formation of baryonic clusters and significantly improves the description of minimum bias data, particularly in the high-multiplicity region and with respect to flavour observables. Following this, we investigate the structure of colour reconnection from a perturbative point of view by considering soft gluon exchanges. We develop a Monte Carlo program, which allows us to evolve the colour structure of systems consisting of up to five clusters and study the resulting colour structures numerically. Next, we investigate the production mechanism of strange quarks in the hadronization model of Herwig 7. We introduce a novel mechanism for strange quark production at different stages of hadronization which depends on the immediate environment to make the process more dynamic and collective. Finally, we improve the multiple parton interaction model of Herwig 7. We tune the resulting model to minimum bias and underlying event data covering a centre-of-mass energy from 200 GeV to 13 TeV and find that we are able to describe all considered observables with one set of parameters.

Zusammenfassung

Nicht-störungstheoretische Modelle sind ein wesentlicher Bestandteil der Simulation von hadronischen Hochenergie-Kollisionen. In dieser Arbeit untersuchen, implementieren und verbessern wir mehrere nicht-störungstheoretische Modelle des Monte-Carlo-Ereignisgenerators Herwig 7. Wir entwickeln zunächst ein neues Modell für die Farbwiederverbindung, welches die Bildung von baryonischen Clustern ermöglicht und die Beschreibung von Daten mit minimalen Detektionskriterien, insbesondere in Regionen hoher Teilchenvielzahl und in Bezug auf die beobachtbare Teilchenart verbessert. Anschließend untersuchen wir die Struktur der Farbwiederverbindung unter störungstheoretischen Gesichtspunkten indem wir den Austausch weicher Gluonen studieren. Wir entwickeln ein Monte-Carlo Programm, welches uns erlaubt die Farbstruktur von Systemen, die aus bis zu fünf Clustern bestehen, zu entwickeln und die resultierenden Farbstrukturen numerisch zu untersuchen. Als nächstes studieren wir den Produktionsmechanismus seltsamer Quarks des Hadronisierungsmodells von Herwig 7. Um den Prozess dynamischer und kollektiver zu gestalten, führen wir an verschiedenen Stellen der Hadronisierung einen neuartigen Mechanismus für die Produktion seltsamer Quarks ein, der von der unmittelbaren Umgebung abhängig ist. Abschließend verbessern wir das Mehrteilcheninteraktionsmodell von Herwig 7. Wir bestimmen die freien Parameter des Modells anhand verfügbarer Daten aus Messungen mit minimalen Detektionskriterien und Messungen der hintergrundartigen Ereignisse, welche eine Massenschwerpunktsenergie von 200 GeV bis 13 TeV umfassen und stellen fest, dass es uns möglich ist alle betrachteten Observablen mit nur einem Satz von Parametern zu beschreiben.

Contents

CHAPTER 1

Introduction

One of the main building blocks of the Standard Model of particle physics is Quantum Chromodynamics (QCD). QCD is a quantum field theory that describes the strong interactions between quarks and gluons and consequently the interactions of composite particles, the hadrons. The theory of QCD has shown to be successful in describing interactions involving large momentum transfer and has been thoroughly tested at the Large Hadron Collider (LHC) by colliding hadrons at energies of up to 13 TeV. In the region of large momentum transfer, the running of the strong coupling α_s is small and perturbation theory can be applied in the context of QCD (perturbative QCD).

However, many aspects of a hadronic collision only involve a small momentum transfer where α_s becomes large and perturbative QCD is no longer applicable. Hence, we have to rely on non-perturbative models for these parts of the simulation to account for a sensible description of data. In this regard, Monte Carlo event generators such as Pythia [1], Sherpa [2] and Herwig [3, 4] play a crucial role. They combine perturbative calculations with non-perturbative models in an attempt to simulate all known aspects of a particle collision and predict exclusive hadronic final states, which ultimately can be compared to data at the particle level taken by collider experiments. A thorough understanding and accurate modelling of the non-perturbative effects contributes significantly to the quality with which we can describe data and make sensible predictions. These non-perturbative models need to be continually reviewed, refined, improved and tested against real data to make progress.

One example for an aspect of a hadronic collision, where non-perturbative modelling is important, is the simulation of the transition from the quarks and gluons originating from a hard interaction to the observed hadrons in the final state. This is an entirely non-perturbative process and we have to rely on so-called hadronization models to describe it. The two most commonly used hadronization models are the string hadronization model [5] and the cluster hadronization model [6]. Both models are based on theoretical considerations and heavily rely on free parameters which have to be determined with fits to experimental data.

Another example where non-perturbative models become important deals with the occurrence of multiple parton interactions. These interactions typically happen at smaller momentum transfer than the main hard interaction and continue to very low scales, where perturbation theory starts to break down. Multiple parton interactions significantly contribute to the underlying event (UE), which contains all hadronic activity in a hadronic collision that is not related to the main hard scattering process, and dominate minimum bias (MB) measurements, where the least number of possible experimental cuts are imposed. A good simulation of multiple parton interactions is also necessary for e.g. the study of jet observables and the searches for new physics since particles from all production processes can contribute to a given final state.

Furthermore, the complex and dense environment of hadronic collisions composes additional challenges. In the QCD colour algebra, the colour structure of an event can be represented with colour lines connecting the different partons participating in the interaction. The colour connections between the partons define the initial conditions for hadronization models, and as such, affect the distribution of particles in the final state. From first principles it is unclear how to calculate the colour connections between different scattering centres of multiple parton interactions. It is expected that further dynamics, which are responsible for additional colour exchanges, lead to a modification of the colour structure and therefore a different final state. In this regard, non-perturbative colour reconnection models are used to establish sensible colour connections based on phenomenological considerations and achieve a good description of the hadronic final state.

All the non-perturbative aspects of the event simulation could contribute significantly to the description of the background, which plays an essential role in the search for new physics, novel phenomena and rare events. Also, the simulation of the pile-up, which resembles the interactions studied in MB analyses, depends on an accurate modelling of non-perturbative effects contributing to the hadronic final state.

This thesis deals with the improvement of the non-perturbative aspects of the simulation of hadronic collisions with the Monte Carlo event generator Herwig 7, and the study of the fundamental aspects of colour reconnection. This thesis is outlined as follows.

Outline of this thesis

Chapter 2 we present an overview about the main concepts of QCD relevant for this work.

Chapter 3 we describe the relevant building blocks to simulate hadronic collisions in the Monte Carlo event generator Herwig 7.

Chapter 4 we present an improved model for colour reconnection. By allowing colour reconnection to form clusters of a baryonic type we significantly improve the description of general MB and flavour observables. This model is also able to describe the enhanced production of strange baryons in high-multiplicity events as seen by ALICE.

Chapter 5 we present our perturbative approach to colour reconnection. Starting from a general QCD amplitude we evolve the colour structure of the amplitude by

means of soft gluon exchanges. We further calculate reconnection probabilities between the original colour structure and the evolved colour structures. We show that our approach leads to properties attributed to colour reconnection and find strong support for geometrical colour reconnection models.

Chapter 6 we review the current strangeness production mechanisms in the cluster hadronization model of Herwig 7 and introduce a new mechanism to produce strangeness at different stages of the event generation. This approach results in a more dynamic production of strangeness which depends on the topology of the event and results in an improved description of data.

Chapter 7 we review the model for the simulation of multiple parton interactions in Herwig 7 and incorporate the simulation of diffractive processes in a more sensible way. We tune the model to data covering an energy range between $\sqrt{s} = 200$ GeV and $\sqrt{s} = 13$ TeV and account for an energy independent description of the model parameters.

Chapter 8 in the last chapter we summarize our results, draw final conclusions and show some avenues for possible future projects.

The model for colour reconnection was originally presented in Ref. [7]. The theoretical approach to colour reconnection was presented in Ref. [8] and the study regarding the production of strangeness based on kinematical considerations was published in Ref. [9]. Similarities between the structure and content of these works and Ch. 4 to Ch. 6 of this thesis are intentional and reflect the contributions of the author.

CHAPTER 2

Quantum Chromodynamics

This thesis deals with the study and implementation of non-perturbative models in the context of the Monte Carlo event generator Herwig 7 to improve the description of the exclusive final state hadrons, which are manifestations of the dynamics on a more elementary level consisting of quarks and gluons. To embed these non-perturbative models in the overall picture of particle physics, we give a brief overview about the main concepts of the theory, which describes the strong interactions between quarks and gluons, namely Quantum Chromodynamics (QCD).

QCD is a non-abelian gauge theory where the underlying group is the group of colour $SU(N_c)$. The QCD Lagrangian reads [10]

$$\mathcal{L}_{\text{QCD}} = \sum_{\Psi=u,d,s,\ldots} \overline{\Psi}_i (i\gamma^\mu D_{\mu,ij} - m_\Psi \delta_{ij}) \Psi_j - \frac{1}{4} F^{a\mu\nu} F^a_{\mu\nu}, \tag{2.1}$$

where the sum runs over the possible quark flavours with corresponding mass m_Ψ. γ^μ denotes the Dirac gamma matrices. The lower case indices i and j run over the number of colours N_c and the upper indices a run from 1 to $N_c^2 - 1$. The covariant derivative D_μ is given by

$$D_{\mu,ij} = \partial_\mu \delta_{ij} - i g_s T^a_{ij} G^a_\mu, \tag{2.2}$$

where g_s is the coupling of QCD and T^a_{ij} are the generators of the $SU(N_c)$ group. The field strength tensor $F^a_{\mu\nu}$ is given by

$$F^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu + g_s f^{abc} G^b_\mu G^c_\nu, \tag{2.3}$$

where G^a_μ is the gluon field and f^{abc} are the structure constants of $SU(N_c)$ defined by the commutation relation of the $SU(N_c)$ group generators

$$\left[T^a, T^b\right] = i f^{abc} T^c. \tag{2.4}$$

The first term in Eq. (2.1) describes the dynamics of quarks and their interactions with gluons. The second term in Eq. (2.1) contains the dynamics of gluons and also the self-interactions of the gluon.

The coupling constant of QCD, g_s, is related to the strong coupling through

$$\alpha_s = g_s^2/4\pi. \tag{2.5}$$

The coupling is not fixed but changes as a function of the energy of the interaction involved. At first order in perturbation theory, the running of the strong coupling can be written as

$$\alpha_s(Q^2) = \frac{1}{b_0 \ln \frac{Q^2}{\Lambda^2}}, \tag{2.6}$$

where $b_0 = (33 - 2n_f)/12\pi$ and n_f represents the number of active quark flavours. Q^2 is the scale of the momentum at which the strong coupling is evaluated. In Eq. (2.6), Λ^2 is a cutoff scale. For large scales, $Q^2 \gg \Lambda^2$, which corresponds to short distances being probed, the strength of the coupling becomes weak and the quarks and gluons can be considered as free particles. This feature of QCD is called *asymptotic freedom* and allows us to use perturbation theory to calculate QCD processes [11, 12].

If the scale of the interaction decreases, the strength of the coupling increases until a point where perturbation theory is no longer applicable. At a very low energy scale $Q^2 \sim \Lambda^2$ the quarks and gluons can no longer be considered free particles and are confined into hadrons. This phenomenon is referred to as *colour confinement*, which means that no individual coloured partons can be observed in nature. The scale at which colour confinement occurs is referred to as the hadronization scale $\Lambda^2 \equiv \Lambda^2_{\text{QCD}}$. At hadron colliders such as the LHC, QCD can be directly applied to calculate the fundamental interactions between the constituents of the colliding hadrons, the quarks and gluons. In Fig. 2.1, we show a sketch of the interaction between two colliding protons. This figure illustrates the necessity for the perturbative and non-perturbative concepts to describe the exclusive final state of a hadronic collision. The grey rectangular area in the center represents the main hard interaction, the part of the collision which happens at high energy scales (small distances). In this region perturbation theory can be used to calculate the process with the methods of perturbative QCD. However, the incoming protons and the description of the exclusive final state constitute both to the non-perturbative part of the collision. While the content of the protons can be parametrized with so-called parton distribution functions (PDFs), we have to rely on models to evolve the physics from the scale of the hard interaction to the non-perturbative scale of the final state particles.

In order to calculate hadronic cross-sections in the parton model with QCD, the factorization theorem can be used [13–15]. For example, the hadronic cross-section for a $pp \to q\bar{q}$ process mediated through the interaction of two gluons reads,

$$\sigma_{\text{had}} = \int \mathrm{d}x_g \mathrm{d}x_g f_g^{h_1}(x_a, Q^2) f_g^{h_2}(x_b, Q^2) \times \hat{\sigma}_{gg\to q\bar{q}}(x_1, x_2, Q^2) \times \text{Final State}, \tag{2.7}$$

where $f_g^h(x, Q^2)$ is the PDF, which contains the probability distribution of finding the gluon with momentum fraction x inside the parent hadron at scale Q^2. $\hat{\sigma}_{gg\to q\bar{q}}$ represents the perturbative cross section of the process in question, the part of the interaction

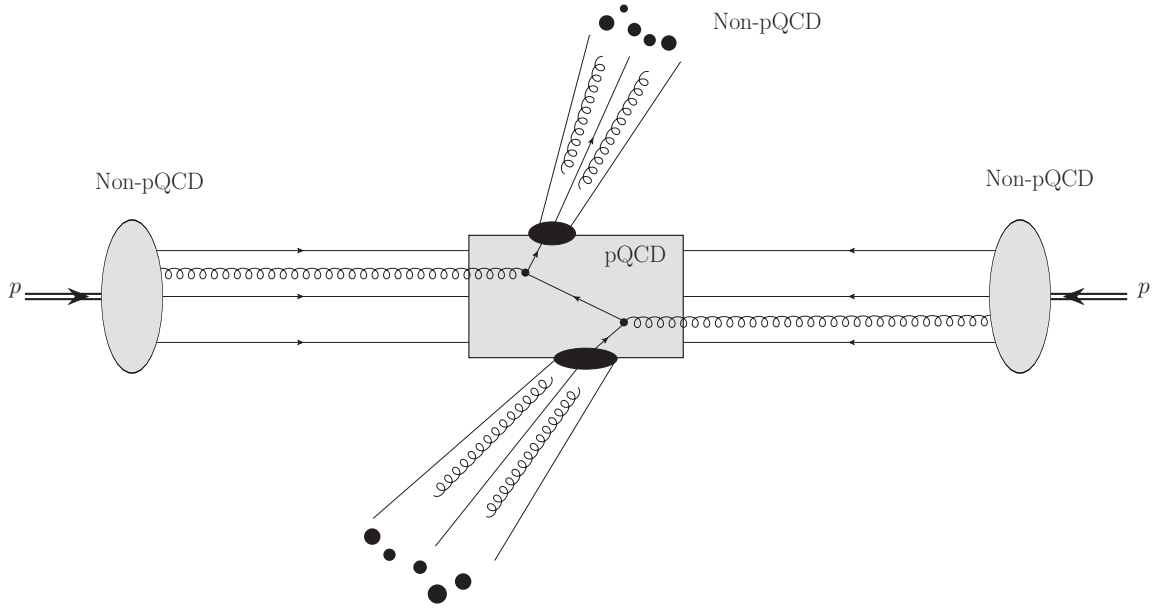

Figure 2.1: Simplified sketch of a proton-proton collision. Shown are the two incoming protons, the three valence quarks of the protons and a very common QCD $2 \to 2$ process ($gg \to q\bar{q}$), namely the hard interaction between two gluons leading to two collimated jets in the final state. The regions where perturbative QCD is applicable is indicated by pQCD and the regions where non-perturbative approaches and models are necessary are indicated by Non-pQCD.

which can be calculated with perturbation theory, and *Final State* describes the evolution of the outgoing partons into the observed final state hadrons. The hadronic cross section for this process can be evaluated by integrating over all possible initial states (PDF) and all possible final states, which depend on the modelling of the transition from the partons to the final state hadrons. The calculation can be made less sensitive to the exact form of the final state modelling by making it more inclusive. For example, if we are not interested in the flavour composition of the final state and only consider the cross section for the production of 2 jets $\sigma_{\text{had}} \equiv \sigma_{jj}$, the calculation becomes less sensitive to the intricacies of the final state modelling. While this approach is very suitable for theoretical calculations, it does not represent the physical reality. Collider experiments are, by its very nature, performing non-inclusive measurements and are therefore sensitive to the physics involved in the modelling of the non-perturbative final state. Furthermore, experimental analyses apply several restrictions on the observed final state, ranging from simple transverse momentum and rapidity cuts to more exclusive cuts like flavour composition.

In order to achieve a sensible description of exclusive measurements, all aspects of Eq. (2.7) have to be understood, this also includes the modelling of the final state. In order to simulate the outcome of hadronic collisions, we use Monte Carlo event generators. We will explain the different stages necessary for the simulation of a hadronic collision with the Monte Carlo event generator Herwig 7 in Ch. 3. Besides being essential for experimental and theoretical studies, they also allow us to test various modifications to the modelling of the final state and then directly compare with data from measurements (see Ch. 6).

Much of this thesis deals with the colour structure and the change in colour structure of QCD processes. Here we quickly recap the basics necessary for the understanding of later chapters. The colour structure in a perturbative QCD process evolves through the interaction of quarks and gluons. The Feynman rule for the quark-gluon vertex is given by

$$-ig_s\gamma^\alpha T^a_{ik}, \tag{2.8}$$

where the term relevant for the colour structure is the $SU(N_c)$ group generator T^a_{ik}. The other term γ^α contains the kinematic information. The colour structure of a gluon exchange between two quark lines is given by

$$T^a_{ik}T^a_{jl}, \tag{2.9}$$

which can be rewritten with the Fierz identity as

$$T^a_{ik}T^a_{jl} = \frac{1}{2}\left[\delta_{il}\delta_{jk} - \frac{1}{N_c}\delta_{ik}\delta_{jl}\right], \tag{2.10}$$

where the Kronecker deltas represent how colour is flowing along the quark lines. Note that the first term corresponds to a colour exchange and results in a change in colour structure for the scattered particles. The second term, which is suppressed by a factor of $1/N_c$ corresponds to a singlet exchange and has no effect on the colour structure of the process.

In QCD calculations, dealing with the full colour structure of a scattering process is often not feasible. To reduce the complexity from the calculation, the $N_c \to \infty$ limit, also known as the leading-colour approximation, is often used [16]. In this limit, Eq. (2.10) simplifies to

$$T^a_{ik}T^a_{jl} \sim \frac{1}{2}\delta_{il}\delta_{jk}, \tag{2.11}$$

and the second term can be neglected. In the $N_c \to \infty$ limit, a gluon exchange corresponds to colour exchange between the two quark lines. Much of ongoing work deals with the incorporation of sub-leading colour corrections in Monte Carlo event generators. One possibility to incorporate sub-leading colour effects is with so called *colour reconnection* models. In Ch. 4, we study a possible extension to the existing colour reconnection model in Herwig 7 and in Ch. 5, we study the connection between a novel perturbative approach to colour reconnection and the cluster hadronization model.

CHAPTER 3

Monte Carlo Event Generators

Monte Carlo event generators are commonly used tools in high-energy physics to simulate particle collisions and predict the exclusive hadronic final states, which can be compared to data taken at collider experiments. The foundation of the simulation is the factorization theorem, which allows us to break the evolution of a hadronic collision down into different steps characterized by the relevant scale of momentum transfer in each part of the subsequent evolution.

While the description of the hard scattering process is well defined within perturbative QCD, we have to rely on phenomenological models to describe the transition from the unstable partons to the observed final state hadrons. Here we give an overview of the event generation with the multi-purpose Monte Carlo event generator Herwig 7, with a focus on the non-perturbative aspects of event generation. Namely, hadronization, colour reconnection, multiple parton interactions and diffraction. We start our discussion with a brief overview of the main building blocks necessary for the simulation of a high-energy particle collision [17]:

- **Hard scattering process:** in the hard scattering process, the incoming partons interact at energy scale Q^2 to produce the primary outgoing partons. The partons and their momenta are generated according to the relevant matrix element for the process. The hard scattering process functions as the starting point of the event simulation by determining the relevant scales and colour-flow information defining the initial conditions for the next stage of the simulation, the parton shower.
- **Parton shower:** the parton shower evolves the coloured partons emerging from the hard interaction at the scale Q^2 down to the low energy scale of the infrared cutoff Q_0^2 by subsequent radiation of other coloured partons. The produced partons also undergo showering themselves. Below the infrared cutoff scale, perturbation theory is no longer applicable and hadronization models are necessary for the next step of the simulation.

- **Hadronization:** the hadronization model describes the evolution of the coloured partons after the parton shower terminates. It converts the colour connected partons into colourless objects, so-called *clusters* and models their decay into unstable hadrons.
- **Decays:** the unstable hadrons emerging from the hadronization model need to be further decayed into the final state hadrons. The possible decay modes are selected according to the experimentally measured branching ratios.

These steps are sufficient for the simulation of a leptonic collisions, for example $e^+e^- \to q\bar{q}$ at the Large Electron-Positron Collider (LEP). To describe hadronic collisions, additional effects need to be taken into account:

- **Multiple parton interactions:** since hadrons are composite objects, the incoming partons that are not involved in the primary hard scattering process can interact with each other as well. Depending on the transverse momentum of the collision, we distinguish between semi-hard interactions and soft interactions.
- **Diffraction:** in order to describe the full scope of possible events, the simulation of diffractive processes, which are not described within the framework of factorization, needs to be included. In Herwig 7, this is done via a simple matrix element for single and double diffractive processes where the dissociated protons are treated fully non-perturbatively.
- **Colour reconnection:** when several partonic interactions occur, it is unclear how to determine the colour connections between the partons emerging from the different scattering centres. Colour reconnection models are then used to rearrange the colour-flow of the event before the hadronization stage and arrive at a sensible description of data.

In the following sections we describe the mentioned building blocks in more detail. We start with the hard process in Sec. 3.1, followed by an explanation of the parton shower in Sec. 3.2. After we discuss the cluster hadronization model in Sec. 3.3 and decays in Sec. 3.4, we describe the model for the simulation of multiple parton interactions in Sec. 3.5 and the simulation of diffractive processes in Sec. 3.6. We end our discussion of Herwig 7 with a review of colour reconnection in Sec. 3.7.

3.1 The Hard Process

A vital ingredient of an event generator is the simulation of the hard scattering process. For hadronic collisions, the cross-section for a partonic $2 \to n$ process can be written as

$$\begin{aligned}\sigma_{h_1h_2\to n} = &\sum_{a,b}\int \mathrm{d}x_a \mathrm{d}x_b \int \mathrm{d}\Phi_n f_a^{h_1}(x_a,Q^2) f_b^{h_2}(x_b,Q^2) \\ &\times \frac{1}{2\hat{s}}|\mathcal{M}_{ab\to n}|^2(\Phi_n;Q^2),\end{aligned} \tag{3.1}$$

where $f_i^h(x,Q^2)$ are the PDFs for finding parton i in hadron h with momentum fraction x at scale Q^2. $\mathrm{d}\Phi_n$ is the Lorentz invariant phase space element and $\mathcal{M}_{ab\to n}$ are the amplitudes of the partonic processes that can be calculated perturbatively.

While Herwig 7 has some built-in PDFs, it is possible to use the PDFs provided by LHAPDF [18]. For simple processes Herwig 7 has built-in matrix elements at LO and NLO in perturbative QCD. With an interface to external matrix element providers like MadGraph [19] or VBFNLO [20], it is possible to include more complicated matrix elements. The format for data exchange between different programs is specified by the *Les Houches Accord* [21].

The Matchbox module [22], which is a major feature of Herwig 7, makes it possible to assemble full NLO QCD matched or merged cross-sections. The Matchbox module can use tree-level, and one-loop matrix elements from external matrix element providers to construct LO and NLO cross-section calculations.

Based on Eq. (3.1), the flavour, the kinematics and the colour-flow of the external particles are sampled via Monte Carlo methods. The colour-flow information and the scales determined by the hard process are then used as the initial conditions for the next step in the event generation workflow: the simulation of the parton shower.

3.2 Parton Showers

Parton showers evolve the partons from the scale of the hard interaction Q^2 down to the infrared cutoff scale of order $Q_0^2 \sim 1\,\mathrm{GeV}$ through the radiation of gluons and the subsequent splitting of gluons into more gluons or quark-antiquark pairs. In that sense, the evolution process approximates the effects of higher-order corrections of perturbative QCD. While being exact in the collinear limit, the approximation also gives reasonable results in the entire phase space. We additionally need to consider the showering of particles before the hard interaction takes place, which is referred to as initial state parton shower. The initial state parton shower evolves the partons of the hard interactions via backwards evolution [23] from the scale of the hard interaction to the low scale of the external hadrons where the PDFs are used to guide the evolution. The partons emitted in the initial state shower proceed to produce final state showers as well. The parton shower is formulated in the $N_c \to \infty$ limit, where N_c is the number of colours. In this limit the gluons can be represented by carrying colour and anticolour. Ongoing work focuses around the inclusion of subleading colour corrections to parton showers [24–27]. Another possibility to account for subleading colour effects is *colour reconnection* (see Sec. 3.7), which changes the colour structure of an event, and is applied after the parton shower terminates. Herwig 7 contains two different parton showers: the angular-ordered $\tilde{q}$-shower [28] and the dipole shower [22, 29]. In this work, we exclusively use the angular-ordered parton shower but for completeness also briefly summarize the dipole shower.

3.2.1 Angular-ordered shower

Herwig 7 uses the coherent branching algorithm [28] to simulate the angular-ordered parton shower. Besides determining the kinematics of the branching, the parton shower decides how colour flows from one leg to another during a branching. In order to treat the colour information of the gluons during the parton shower evolution, there are different possibilities. Based on the choice, it is possible to reduce or enhance the amount of additional radiation. The standard when using NLO corrections or external LHE files is to choose the colour connected evolution partner of the gluon randomly.

The splitting kernels used in the angular-ordered parton shower are the quasi-collinear Altarelli-Parisi splitting functions [30–32]. An important property of the angular-ordered parton shower is the so-called *pre-confinement* property of QCD [33,34], which means that for leptonic collisions, at each point in the evolution, the parton shower forms colour connected singlet combinations of partons with an asymptotically universal invariant mass distribution. Universal means that the invariant mass distribution only depends on the cutoff scale of the parton shower and not on the centre-of-mass energy of the collision or the hard process itself [35].

3.2.2 Dipole shower

In the $N_c \to \infty$ limit, the colour-flow of an event can be decomposed into a set of planar colour lines. Every colour line connects a coloured parton with an anticoloured antiparton and forms a colour-anticolour dipole. The parton shower then generates radiation according to the dipole radiation pattern of a pair of colour connected partons. When a dipole emits a gluon, the dipole splits into two new dipoles. The dipoles proceed to split until they reach the infrared cutoff Q_0^2, which terminates the shower. Since each emission with a finite transverse momentum results in a recoil of the system, the recoil may affect the subsequent evolution of neighbouring dipoles.

The dipole shower generation follows an ordering of the emission in which the emission of the parton with the highest transverse momentum from any dipole is generated first. This transverse momentum is the upper limit for the subsequent evolution of dipoles. The basis of the dipole shower in Herwig 7 is the Catani-Seymour subtraction kernel [36].

3.3 Hadronization

After the parton shower has evolved the partons originating from the hard interaction down to the scale of the infrared cutoff $Q_0^2 \sim 1\,\mathrm{GeV}$, the applicability of perturbative QCD starts to break down. In order to describe the transition from the partons to the final state hadrons, non-perturbative hadronization models are required. Herwig 7 uses the cluster hadronization model [37], which is based on the pre-confinement property of the angular-ordered parton shower [33, 34]. As a consequence, the colour connected partons that form colour singlets should be close in momentum-space. In the $N_c \to \infty$ limit, the colour-flow of a QCD process is representable as a set of planar colour lines [16]. The gluons can be represented as carrying colour and anticolour and are connected to the colour lines of two other partons. The hadronization model starts by splitting the remaining gluons at the end of the parton shower into quark-antiquark pairs. This splitting results in a partonic final state that is only composed of colour-connected quarks and antiquarks. These colour-connected quark-antiquark pairs form colour singlets, the so-called *clusters* and represent excited hadronic states which further decay into the observed final state hadrons. The hadronization model of Herwig 7 consists of several different stages, here given in algorithmically occurring order:

- Non-perturbative gluon splitting,
- Colour reconnection,
- Cluster fissioning,

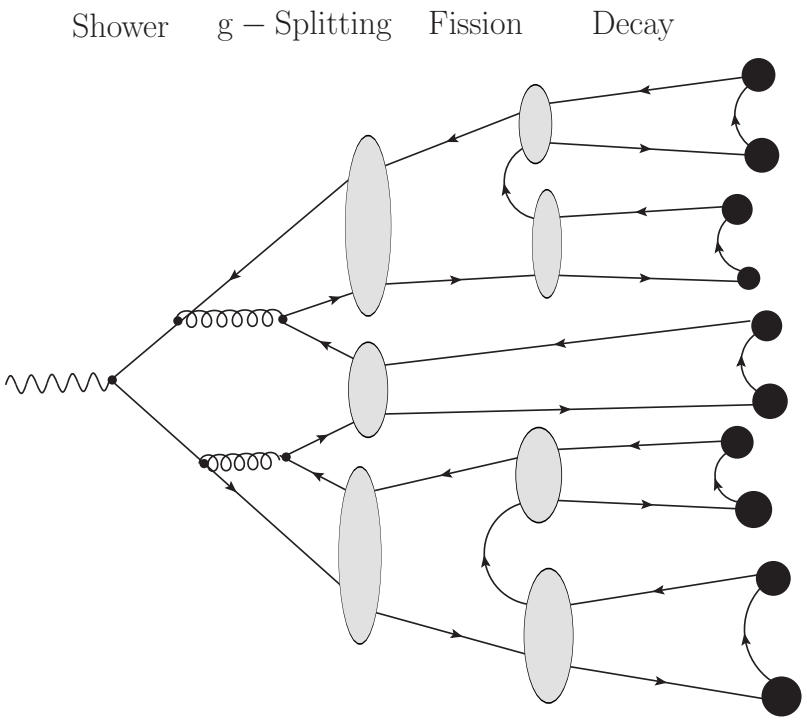

Figure 3.1: Sketch of a simplified event where we show the major stages of hadronization after the parton shower: gluon splitting, cluster fissioning and the cluster decay. Grey ellipses are clusters, while black circles represent hadrons.

- Cluster decay to hadrons,
- Unstable hadron decays.

We show a sketch of the different stages of hadronization in Fig. 3.1, where we have omitted the colour reconnection stage for simplicity. From a technical point of view colour reconnection is part of the hadronization model of Herwig 7 but since much of this thesis is related to work on colour reconnection, a discussion of colour reconnection will be given in Sec. 3.7.

3.3.1 Non-perturbative gluon splitting

After the parton shower evolution, the remaining gluons are split non-perturbatively into quark-antiquark pairs where only up or down quarks are allowed. The flavour of the quarks is determined by a weight which is a tunable parameter in the hadronization model. The gluon needs to have at least twice the mass of the quark it splits into $m_g > 2m_q$, where the gluon mass is a free, non-perturbative parameter of the cluster hadronization model. The splitting $g \to q\bar{q}$ is done isotropically in the rest frame of the gluon. Once all remaining gluons have split, the event only consists of colour-connected quark-antiquark pairs. These colour connected quark-antiquark pairs form colourless singlets, the clusters, where the resulting cluster momentum is the sum of the constituent momenta. The properties of a cluster depend entirely on the flavour of its constituent quarks and the invariant cluster mass

$$M^2 = (p_q + p_{\bar{q}})^2, \tag{3.2}$$

where $p_{q,\bar{q}}$ are the four-momenta of the constituent quarks. The invariant cluster mass determines whether a cluster needs to undergo cluster fission or can directly decay into hadrons. The distribution of invariant cluster masses peaks at low energies, which makes the interpretation as excited hadronic states physically sensible.

3.3.2 Cluster fissioning

A small fraction of the clusters consists of colour connected quark-antiquark pairs which are not close in momentum space and are therefore too heavy to decay into hadrons directly. These heavy clusters are allowed to fission into two lighter, less excited clusters if the invariant mass M of the original cluster satisfies

$$M^p \geq \mu^p + (m_1 + m_2)^p, \tag{3.3}$$

where m_1, m_2 are the masses of the constituent quarks. p and μ are free parameters of the model, which have different values for different quark flavours. After the selection of clusters, which need to undergo fission based on Eq. (3.3), a $q\bar{q}$ pair is created where the flavour of the quarks is drawn from light quarks (u, d, s) with a fixed weight. Once the quark flavour is selected the original cluster fissions into two new clusters with one of the original quarks in each new cluster. The invariant masses of the two new clusters M_i are determined according to the distribution

$$M_i = m_i + (M - m_i - m_q)\mathcal{R}_i^{1/w}, \qquad i = 1, 2, \tag{3.4}$$

where m_i are the masses of the constituent quarks of the original cluster, M is the invariant mass of the original cluster, m_q is the mass of the newly created quark, and w is a parameter that controls the shape of the distribution. The parameter w has different values for clusters containing light quarks, charm or bottom quarks. $\mathcal{R}$ is a random number between 0 and 1.

The fission into two new clusters is accepted if the sum of the new invariant cluster masses is less than the invariant mass of the original cluster $M_1 + M_2 \leq M$ and the sum of the constituent masses of the new cluster is less than the invariant mass of the new cluster $m_{q_i} + m_{\bar{q}_i} \leq M_i$.

A different approach is taken when one or both constituents of the cluster are part of the beam remnant. The remnant is the remainder of the hadron after the partons participating in the hard interaction are removed. In order to suppress hadronic activity from clusters originating from beam remnants, the invariant cluster masses of the two new clusters are determined according to the distribution

$$M_i = m_i + m_q + x, \tag{3.5}$$

where x is sampled between 0 and $M - m_1 - m_2 - 2m_q$ from a distribution which falls off exponentially according to

$$\frac{\mathrm{d}P}{\mathrm{d}x^2} = e^{-bx}, \tag{3.6}$$

where b is a tunable parameter and determines the slope of the distribution. This way, the clusters produced from the remnant clusters are in general light, fast and do not contribute much to the overall hadronic activity.

3.3.3 Cluster dacays

After the fission of heavy clusters, the clusters decay into excited hadrons. Given a cluster with constituents $q_1, \bar{q}_2$, the weight for producing hadrons $h_a = q_1\bar{q}$, $h_b = q\bar{q}_2$, with q ($\bar{q}$) denoting a quark (antiquark) or diquark (antidiquark) is given by

$$\mathcal{W}(h_a, h_b) = P_q w_a s_a w_b s_b p^*_{a,b}, \tag{3.7}$$

where P_q is the production weight for the given quark or diquark species, w_i are the weights for the relevant hadron production and s_i are the suppression factors of the corresponding hadrons. The remaining factor in the weight p^* is the available two-body phase space for the decay. After the decay products are determined, the cluster decays isotropically in its rest frame. If a cluster contains a quark from the perturbative stage of the event, the resulting hadron is smeared along the direction of that parton where the polar angle is sampled via

$$\cos\Theta_{\text{smr}} = 1 + \rho \log \mathcal{R}\ , \quad \text{s.t.} \quad \cos\Theta_{\text{smr}} > -1, \tag{3.8}$$

where ρ is a free parameter and has different values for different quark flavours. The azimuthal angle is distributed uniformly between 0 and 2π.

If a cluster is too light to decay into two new hadrons, it decays into the lightest allowed hadron instead. In order to give the hadron the correct physical mass, energy and momentum are reshuffled between neighbouring clusters.

3.4 Decays

In the final stage of hadronization we need to describe the decay of unstable hadrons into stable hadrons. These decays are simulated in Herwig 7 with the model for hadronic decays described in Refs. [38, 39]. Herwig 7 uses up-to-date particle properties, such as masses, widths, lifetimes, decay modes and branching ratios from the Particle Data Group's (PDG) compilation [40]. A full treatment of spin correlation effects is taken into account using the algorithms of Refs. [41–44] for the decays of all unstable particles. Various matrix elements for hadron and tau decays including matrix elements for the spin structure of the decays and specific matrix elements for important decay modes are implemented in Herwig 7. The branching ratios and properties for mesons containing b quarks, $B_{u,d,s}$, are taken from the data tables of EvtGEN [45], which was tuned to B-factory data. It is furthermore possible to use EvtGEN directly for the simulation of B decays.

3.5 Multiple Parton Interactions

A single hadron scattering event is expected to give rise to multiple interactions between the constituents of the hadrons [46]. The model for the simulation of multiple parton interactions in Herwig 7 is based on the eikonal model from Refs. [47–50]. It consists of a perturbative and a non-perturbative part. The model for soft interactions was introduced in Refs. [51,52] and revised in Ref. [53], where the description of elastic gluon $2 \to 2$ scattering has been replaced with particle production obeying multiperipheral kinematics [54, 55]. This model solved many problems of the old model, including the infamous *bump* problem [56, 57].

The starting point for our discussion about multiple parton interactions is the cross-section for QCD dijet production, which can be calculated according to

$$\sigma_{\text{dijet}}(s; p_\perp^{\min}) = \sum_{i,j} \int \mathrm{d}x_1 \mathrm{d}x_2 \int_{p_\perp^{\min,2}} \mathrm{d}p_\perp^2 f_i^{h_1}(x_1, Q^2) f_j^{h_2}(x_2, Q^2) \frac{\mathrm{d}\sigma_{ij}}{\mathrm{d}p_\perp^2}, \tag{3.9}$$

where $\mathrm{d}\sigma_{ij}$ is the differential partonic cross-section for QCD $2 \to 2$ scattering processes. $p_\perp^{\min}$ is the lower cutoff on the transverse momentum which has to be large enough to ensure perturbative QCD is still applicable. For $p_\perp^{\min} = 2\,\text{GeV}$ the calculated cross-section from Eq. (3.9) exceeds the total cross-section already at $\sqrt{s} \sim 1$ TeV [58]. This violation of unitarity is resolved by interpreting the QCD dijet cross-section as an inclusive cross-section $\sigma_{\text{dijet}} \equiv \sigma_{\text{dijet}}^{\text{inc}}$. The inclusive dijet cross-section exceeds the inelastic cross-section by a factor equal to the average number of dijet production above $p_\perp^{\min}$

$$\sigma_{\text{dijet}}^{\text{inc}}(s, p_\perp^{\min}) = \langle n_{\text{dijet}} \rangle \cdot \sigma_{\text{inel}}(s, p_\perp^{\min}), \tag{3.10}$$

where σ_{inel} is required to be less than the total cross-section. In this way the occurrence of additional QCD interactions unitarises the QCD cross section.

3.5.1 Semi-hard and soft interactions

We distinguish between two types of interactions depending on their transverse momentum: interactions above a certain cutoff value $p_\perp^{\min}$ that are simulated as QCD $2 \to 2$ processes and are referred to as *semi-hard* interactions and interactions below $p_\perp^{\min}$ that are modelled as multiperipheral particle production and are referred to as *soft* interactions.

Following Refs. [48, 50], the average number of semi-hard interactions at a fixed impact parameter $b \equiv |\mathbf{b}|$, is given by

$$\langle n(\mathbf{b}, s) \rangle_{\text{hard}} = A(b, \mu_{\text{hard}}) \sigma_{\text{hard}}^{\text{inc}}(s, p_\perp^{\min}), \tag{3.11}$$

where $A(b, \mu_{\text{hard}})$ describes the overlap of the two colliding hadrons in terms of the impact parameter b and the inverse proton radius μ_{hard}. $\sigma_{\text{hard}}^{\text{inc}}(s, p_\perp^{\min})$ is the inclusive cross-section for QCD $2 \to 2$ processes above $p_\perp^{\min}$ which can be calculated in perturbative QCD according to Eq. (3.9).

In order to describe low-$p_\perp$ jet production, the model was extended to incorporate interactions below $p_\perp^{\min}$ in Ref. [52]. The average number of soft interactions is calculated according to

$$\langle n(\mathbf{b}, s) \rangle_{\text{soft}} = A(b, \mu_{\text{soft}}) \sigma_{\text{soft}}^{\text{inc}}(s), \tag{3.12}$$

where the overlap function between the two colliding hadrons $A(b, \mu_{\text{soft}})$ has the same structure as in Eq. (3.11) but with a free parameter μ_{soft} that allows soft interactions to have a different matter distribution inside the hadron. The soft inclusive cross-section $\sigma_{\text{soft}}^{\text{inc}}(s)$ cannot be calculated from first principles. Instead it is chosen such that the total cross-section $\sigma_{\text{tot}}(s)$ is correctly described. In both cases the overlap function $A(b, \mu)$ must satisfy

$$\int \mathrm{d}^2\mathbf{b}\ A(b, \mu) = 1\ . \tag{3.13}$$

The spatial matter distribution inside the hadrons is assumed to have the same functional form as the electromagnetic form factor of the proton, which results in the following parametrization of the overlap function

$$A(b,\mu) = \frac{\mu^2}{96\pi}(\mu b)^3 K_3(\mu b), \tag{3.14}$$

where K_3 is a modified Bessel function of the second kind. If one further assumes that the probability distributions of semi-hard and soft interactions are independent of each other, and both scattering types are uncorrelated and therefore obey Poissonian statistics, we can write the probability of having h semi-hard and n soft interactions for a given impact parameter as

$$\mathcal{P}_{h,n}(\mathbf{b},s) = \frac{\langle n(\mathbf{b},s)\rangle^h_{\text{hard}}}{h!}\frac{\langle n(\mathbf{b},s)\rangle^n_{\text{soft}}}{n!}e^{-(\langle n(\mathbf{b},s)\rangle_{\text{hard}}+\langle n(\mathbf{b},s)\rangle_{\text{soft}})}. \tag{3.15}$$

The cross-section for h semi-hard and n soft interactions can then be calculated by integrating over the impact parameter

$$\sigma_{h,n}(s) = \int \mathrm{d}^2\mathbf{b}\ \mathcal{P}_{h,n}(\mathbf{b},s). \tag{3.16}$$

The cross-section for inelastic events where at least one event is hard or soft is then given by summing over the individual multiplicities and we obtain

$$\sigma_{\text{inel}}(s) = \int \mathrm{d}^2\mathbf{b} \sum_{h,n\geq 1} \mathcal{P}_{h,n}(\mathbf{b},s) = \int \mathrm{d}^2\mathbf{b}\left[1 - e^{-(\langle n(\mathbf{b},s)\rangle_{\text{hard}}+\langle n(\mathbf{b},s)\rangle_{\text{soft}})}\right]. \tag{3.17}$$

The probability for having exactly h semi-hard and n soft interactions, where $h+n \geq 1$, is then given by

$$P_{h,n}(s) = \frac{\sigma_{h,n}(s)}{\sigma_{\text{inel}}(s)}. \tag{3.18}$$

This equation is the core of the multiple parton interaction model of Herwig 7. It defines a matrix containing the probabilities for the various multiplicities of semi-hard and soft interactions. This matrix is evaluated at the beginning of each run and the multiplicities are drawn at the beginning of each event based on the calculated probabilities.

In order to calculate the probability matrix Eq. (3.18), we need to determine the average number of semi-hard and soft events Eqs. (3.11, 3.12), which depend on the parameters μ_{hard}, μ_{soft}, $p_\perp^{\text{min}}$ and $\sigma^{\text{inc}}_{\text{soft}}(s)$. The model is structured that μ_{hard} and $p_\perp^{\text{min}}$ can be tuned to data and directly determine the average number of semi-hard interactions. Instead of tuning the remaining parameters we use a powerful constraint via the eikonal model. By using the experimentally measured total cross-section as an input parameter, we are able to directly determine the parameter values μ_{soft} and $\sigma^{\text{inc}}_{\text{soft}}(s)$ in order to reproduce the total cross section.

Link to the total cross-section via the eikonal model

The elastic scattering amplitude in impact parameter space $a(\mathbf{b},s)$ can be expressed in terms of a real eikonal function $\chi(\mathbf{b},s)$ as

$$a(\mathbf{b},s) = \frac{1}{2i}\left[e^{-\chi(\mathbf{b},s)} - 1\right]. \tag{3.19}$$

The total cross-section and the elastic cross-section can then be parametrized with the eikonal function as

$$\sigma_{\text{tot}}(s) = 2\int \mathrm{d}^2\mathbf{b}\left[1 - e^{-\chi(\mathbf{b},s)}\right], \qquad \sigma_{\text{el}}(s) = \int \mathrm{d}^2\mathbf{b}\left|1 - e^{-\chi(\mathbf{b},s)}\right|^2. \tag{3.20}$$

The inelastic cross-section can then be calculated according to

$$\sigma_{\text{inel}}(s) = \sigma_{\text{tot}}(s) - \sigma_{\text{el}}(s) = \int \mathrm{d}^2\mathbf{b}\left[1 - e^{-2\chi(\mathbf{b},s)}\right]. \tag{3.21}$$

The elastic slope parameter can also be expressed in this framework and yields

$$b_{\text{el}} = \frac{1}{\sigma_{\text{tot}}}\int \mathrm{d}^2\mathbf{b}\; b^2\;\left[1 - e^{-\chi(\mathbf{b},s)}\right]. \tag{3.22}$$

By comparing the formulas for the inelastic cross-sections, Eq. (3.17) and Eq. (3.21), we see that the two formulas coincide if the eikonal is chosen as

$$\begin{aligned}\chi(\mathbf{b},s) =& \frac{1}{2}\left(\langle n(\mathbf{b},s)\rangle_{\text{hard}} + \langle n(\mathbf{b},s)\rangle_{\text{soft}}\right)\\ =& \frac{1}{2}\left(A(b,\mu_{\text{hard}})\sigma^{\text{inc}}_{\text{hard}}(p_\perp^{\min},s) + A(b,\mu_{\text{soft}})\sigma^{\text{inc}}_{\text{soft}}(s)\right).\end{aligned} \tag{3.23}$$

The total cross-section and the elastic slope parameter are input parameters in this model. The value of the total cross-section is taken from the Donnachie-Landshoff parametrization [59, 60]

$$\sigma_{\text{tot}}(s) = \sigma_{\mathbb{P}}\left(\frac{s}{\text{GeV}^2}\right)^{\epsilon} + \sigma_{\mathbb{R}}\left(\frac{s}{\text{GeV}^2}\right)^{-\eta}, \tag{3.24}$$

where $\sigma_{\mathbb{P}} = 21.7\,\text{mb}$, $\epsilon = 0.0805$, $\sigma_{\mathbb{R}} = 56\,\text{mb}(pp), 98\,\text{mb}(p\bar{p})$ and $\eta = 0.0452$. The value for the elastic slope parameter b_{el} is obtained from fits to elastic scattering data from Ref. [61].

In combination with the tuned parameter values for $p_\perp^{\min}$ and μ_{hard}, the value for the soft inclusive cross-section $\sigma^{\text{inc}}_{\text{soft}}(s)$ and the inverse proton radius μ_{soft} can be determined by solving Eqs. (3.20) and Eq. (3.22) with the eikonal function given in Eq. (3.23). An extensive study on the restriction of the allowed parameter space for $p_\perp^{\min}$ and μ_{hard} for different parametrizations of the total cross-section was performed in Ref. [51].

3.5.2 Implementation

While the semi-hard interactions are created as QCD $2 \to 2$ processes, the exact form of the additional soft interactions below $p_\perp^{\min}$ remains ambiguous. In the current model, the soft interactions are modelled according to multiperipheral parton production. Multiperipheral parton production is a $2 \to N$ process where the resulting N partons are evenly distributed in rapidity. In the following, we will refer to the produced partons obeying multiperipheral kinematics as a parton *ladder*. Every ladder generated consists of a quark-antiquark pair and several gluons.

The simulation of the soft interactions follows after the perturbative semi-hard processes have been generated. The number of soft interactions N_{soft} is sampled at the

beginning of each event from the probability matrix Eq. (3.18) and each soft interaction is generated as a parton ladder obeying multiperipheral kinematics.

The total energy available to simulate the additional soft interactions is given by the remaining energy of the remnants. The generation of ladders continues until N_{soft} ladders are created or until there is no remaining energy in the remnants. The number of partons in the ladder is sampled from a Poissonian distribution with the mean value

$$\langle N \rangle = N_0 \times \ln \frac{(p_{r1} + p_{r2})^2}{m_g^2}, \tag{3.25}$$

where p_{r1} and p_{r2} are the four momenta of the remnants, m_g is the non-perturbative gluon mass and N_0 is a tunable parameter of the model. The model is tuned to MB data from different centre-of-mass energies at $\sqrt{s} = 900\,\text{GeV}$, $\sqrt{s} = 7\,\text{TeV}$ and $\sqrt{s} = 13\,\text{TeV}$. The energy dependence of the resulting parameter values for N_0 is parametrized according to

$$N_0 = \left(\frac{s}{\text{TeV}^2}\right)^{-0.08}, \tag{3.26}$$

such that the model returns reasonable results for different centre-of-mass energies [62]. After the number of particles in the ladder is calculated, the energy fraction of the remnants, necessary to distribute the particles equally in rapidity, is calculated according to

$$x = e^{-\frac{\Delta Y}{2N+1}}, \tag{3.27}$$

where ΔY is the rapidity interval between the two remnants and N is the number of particles generated in the ladder. In order to allow for fluctuations, a Gaussian smearing is employed. The remaining degree of freedom is the transverse momentum of the ladder partons $p_{\perp 1,\ldots,N}$. The transverse momenta are sampled below $p_\perp^{\min}$ from a distribution that is parametrized with a Gaussian distribution [58]

$$\frac{\mathrm{d}\sigma_{\text{soft}}^{\text{inc}}(s)}{\mathrm{d}p_\perp^2} = Ae^{-\beta p_\perp^2}. \tag{3.28}$$

Two constraints fix the parameters A and β. First, the integrated differential inclusive soft cross-section has to match the inclusive soft cross-section which was fixed to describe the total cross-section

$$\int \mathrm{d}p_\perp^2 \frac{\mathrm{d}\sigma_{\text{soft}}^{\text{inc}}(s)}{\mathrm{d}p_\perp^2} \overset{!}{=} \sigma_{\text{soft}}^{\text{inc}}(s). \tag{3.29}$$

Second, the differential transverse momentum distribution has to match the perturbative distribution at $p_\perp = p_\perp^{\min}$ and must be continuous at the matching scale

$$H(s, p_\perp^{\min}) \equiv \left.\frac{\mathrm{d}\sigma_{\text{hard}}^{\text{inc}}(\mathrm{s})}{\mathrm{d}p_\perp^2}\right|_{p_\perp = p_\perp^{\min}} \overset{!}{=} \left.\frac{\mathrm{d}\sigma_{\text{soft}}^{\text{inc}}(s)}{\mathrm{d}p_\perp^2}\right|_{p_\perp = p_\perp^{\min}}. \tag{3.30}$$

These condition are fulfilled by the parametrization

$$\frac{\mathrm{d}\sigma_{\text{soft}}^{\text{inc}}(s)}{\mathrm{d}p_\perp^2} = H(s; p_\perp^{\min}) e^{-\beta(p_\perp^2 - p_\perp^{\min 2})}, \tag{3.31}$$

where the slope of the distribution β must satisfy

$$\frac{e^{\beta p_\perp^{\min 2}} - 1}{\beta} = \frac{\sigma_{\text{soft}}^{\text{inc}}(s)}{H(s; p_\perp^{\min})}. \tag{3.32}$$

The individual transverse momenta of the ladder partons are then sampled from the resulting distribution which is determined at the beginning of each run.

The colour connections are set such that the ladder is in a colour singlet state, which means that the partons in the ladder with the positive maximum rapidity and the negative maximum rapidity need to be quarks and antiquarks. The flavour of the quark-antiquark pair is chosen between up and down quarks only. All remaining ladder partons between the quark-antiquark pair are gluons. The colour lines are then connected such that adjacent partons in rapidity are connected with each other.

The gluons in the ladder are treated fully non-perturbatively and split into quark-antiquark pairs. The quark-antiquarks then proceed to form clusters based on their colour connections as explained in Sec. 3.3. We have illustrated this in Fig. 3.2.

3.6 Diffraction

The model for diffraction was initially presented in Ref. [57] and published in Ref. [53]. The combination of the newly introduced model for soft interactions with a model for diffraction led to a significant improvement in the description of MB and UE observables.

Two types of diffractive processes are implemented in Herwig 7[1]. Single diffractive (SD) events, $p + p \to p + X$, where one of the outgoing protons has been dissociated and double diffractive (DD) events, $p + p \to X_A + X_B$, where both of the outgoing protons dissociate. The dissociated protons are represented through X, X_A, X_B respectively and are modelled as a colour connected quark-diquark pair whose diffractive masses are chosen as described below. The diffractive processes are generated according to the differential cross-sections for single and double diffraction by considering the amplitude for a single pomeron exchange which can be described by Regge theory and the generalized optical theorem [63, 64]. Single and double diffractive events are generated according to the respective differential cross-sections which, for small values of momentum transfer $|t|$ simplify to

$$\frac{\mathrm{d}^2\sigma_{\text{SD}}}{\mathrm{d}M^2 dt} = N \left(\frac{s}{M^2}\right)^{\alpha_{\mathbb{P}}(0)} e^{\left(B_0 + 2\alpha' \ln \frac{s}{M^2}\right)t}, \tag{3.33}$$

and

$$\frac{\mathrm{d}^3\sigma_{\text{DD}}}{\mathrm{d}M_A^2 dM_B^2 dt} = N \left(\frac{s}{M_A^2}\right)^{\alpha_{\mathbb{P}}(0)} \left(\frac{s_0}{M_B^2}\right)^{\alpha_{\mathbb{P}}(0)} \times e^{\left(b + 2\alpha' \ln \frac{s s_0}{M_A^2 M_B^2}\right)t}, \tag{3.34}$$

where M, M_A, M_B are the respective diffractive masses, $\alpha_{\mathbb{P}}(0)$ is the soft pomeron intercept, α' is the soft pomeron slope, and B_0 is the proton pomeron slope. The values for these parameters are taken from fits to elastic scattering data from Ref. [59]. The normalization between single and double diffraction is chosen according to measurements of diffractive cross-sections in Ref. [65].

[1] So far only diffractive processes among protons are implemented.

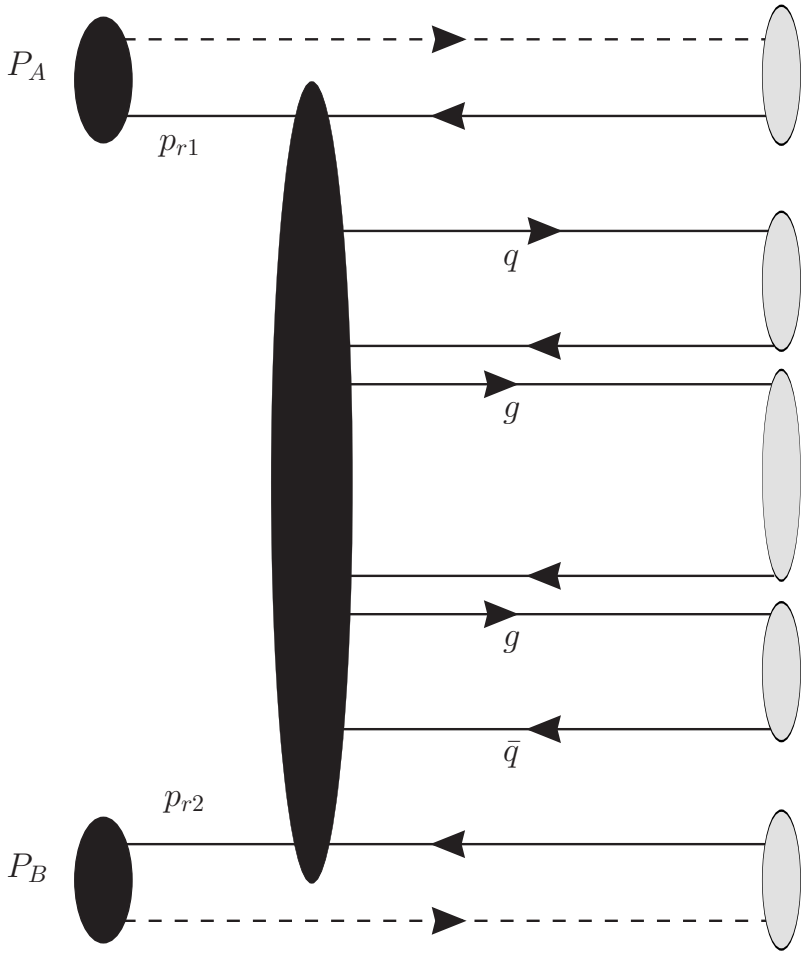

Figure 3.2: Pictorial representation of one soft interaction in Herwig 7 due to the collision of two protons A and B. The soft interaction is modelled as a remnant-remnant interaction (black ellipse spanning over the two remnants). The arrows indicate whether a parton is carrying colour or anticolour. The dashed lines represent quarks which were extracted from the protons. The grey ellipes represent the clusters originating from the soft interaction.

The kinematics are generated by first sampling the momentum transfer t, and the diffractive masses M, M_A, M_B depending on the process in question. The scattering angle θ between the incoming protons and the diffractive systems is then calculated according to

$$\cos\theta = \frac{s\left(s + 2t - 2m_p^2 - M_A^2 - M_B^2\right)}{\lambda(s, M_A^2, M_B^2)\lambda(s, m_p^2, m_p^2)}, \tag{3.35}$$

where

$$\lambda(x, y, z) = x^2 + y^2 + z^2 - 2(xy + yz + xz), \tag{3.36}$$

is the so-called Källén function. If $\alpha_{\mathbb{P}}(0) = 1$ the distributions of diffractive masses of the dissociated protons are approximately given by

$$\left.\frac{\mathrm{d}^2\sigma}{\mathrm{d}M^2\mathrm{d}t}\right|_{t=0} \approx \frac{1}{M^2}. \tag{3.37}$$

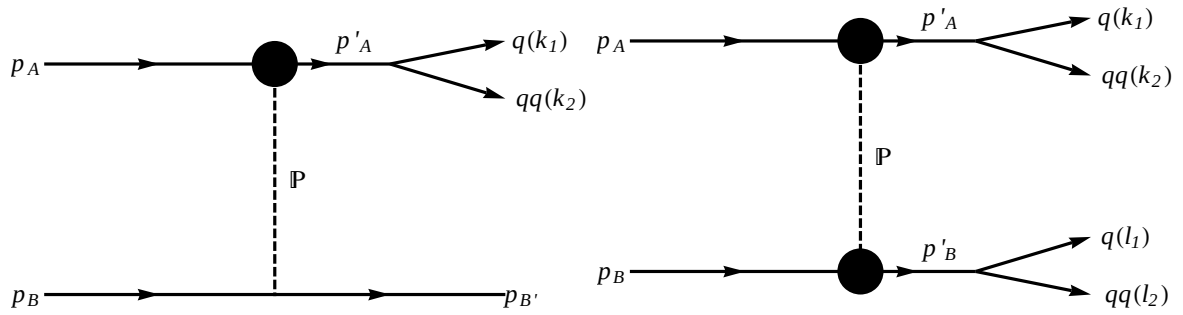

Figure 3.3: Diffractive dissociation for single (left) and double (right) diffraction mediated through pomeron exchange. Figure taken from Ref. [53].

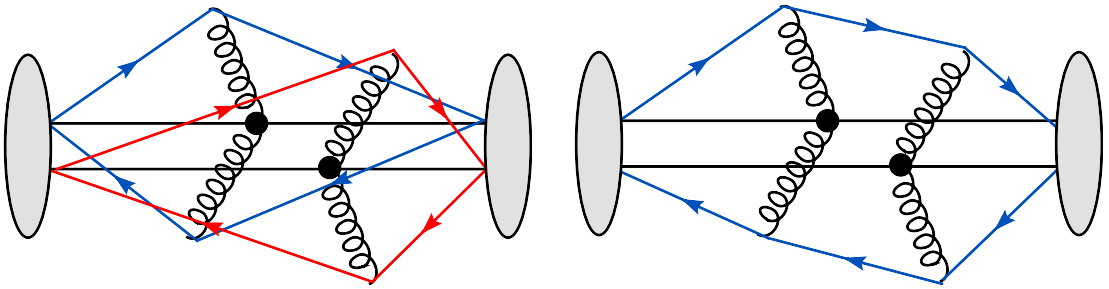

Figure 3.4: Schematic diagram of the colour connections of two multiple parton interactions before colour reconnection (left) and after colour reconnection (right) [66]. In this example, colour reconnection leads to colour connections between partons closer in phase space, which results in lighter clusters.

With the diffractive masses and the scattering angle, the outgoing momenta can be constructed. The dissociated proton is then decayed into a quark-diquark pair moving collinear to the direction of the dissociated proton. The diffractive final state is treated fully non-perturbatively and handled by the hadronization model where the formed clusters decay into the final state hadrons. A pictorial representation of single and double diffractive processes is shown in Fig. 3.3.

3.7 Colour Reconnection

The assignment of colour connections between quark-antiquark pairs is not without flaws. While at e^+e^- collisions the colour connections emerging from the parton shower lead to an asymptotically invariant mass distribution of clusters, the situation becomes more complicated in the presence of multiple parton interactions during hadronic collisions. Since it is unclear how the colour connections between different scattering centres emerge, non-perturbative models are necessary in order to rearrange the colour-flow to arrive at a sensible description of data [46]. A figure of the possible effect of colour reconnection on the colour connections of two multiple parton interactions is shown in Fig. 3.4. In this case the colour reconnection leads to a decrease of charged multiplicity for a given partonic configuration and hence an increase of the transverse momentum per charged particle, as favoured by the relevant observable as shown exemplary in Fig. 3.5. Colour reconnection plays an essential role in the description of various MB observables as well as for the description of high-multiplicity events where extreme event

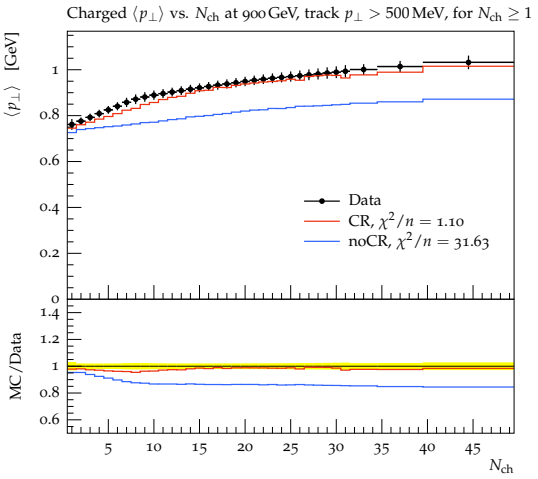

Figure 3.5: Comparison between Herwig 7 with and without colour reconnection to the $\langle p_\perp \rangle$ vs. N_{ch} observable as measured by ATLAS [67]. Figure taken from Ref. [68].

topologies with many overlapping clusters originating from multiple parton interactions are encountered.

A paradigm which inspired at least the development of one model is the so-called notion of a colour pre-confined state, which states that after the evolution of the parton shower has terminated, the colour connected partons are close in momentum space leading to a distribution of invariant cluster masses which peaks at small values of M dictated by the parton shower infrared cutoff. Current developments in this direction, including several space-time hadronization models [69, 70], seem to look promising for future studies.

3.7.1 The colour reconnection models in Herwig 7

Two algorithms for colour reconnection are currently implemented in Herwig 7, the *plain* colour reconnection and the *statistical* colour reconnection [71, 72]. Both algorithms try to find configurations of clusters that would reduce the sum of squared invariant cluster masses

$$\lambda = \sum_{i=1}^{N_{\text{cl}}} M_i^2, \tag{3.38}$$

where N_{cl} is the number of clusters in an event. The plain colour reconnection algorithm picks a cluster randomly from the list of clusters and compares it to all other clusters. For every cluster the invariant masses $M_{\text{A}}^2 + M_{\text{B}}^2$ of the original cluster configuration and the masses $M_{\text{C}}^2 + M_{\text{D}}^2$ of the possible new clusters are calculated. The cluster configuration that results in the lowest sum of squared invariant cluster masses is then accepted for reconnection with a fixed probability p_{R}. If the reconnection is accepted, the clusters (A) and (B) are replaced by the clusters (C) and (D). This algorithm results in clusters of the same initial mass (i.e no reconnection), or less massive clusters. In Fig. 3.6 we show the distribution of invariant cluster masses for a centre-of-mass energy of $\sqrt{s} = 7\,\text{TeV}$ before and after the plain colour reconnection. Due to the colour reconnection procedure, the distribution is enhanced in the low-mass region and suppressed in the high-mass tail.

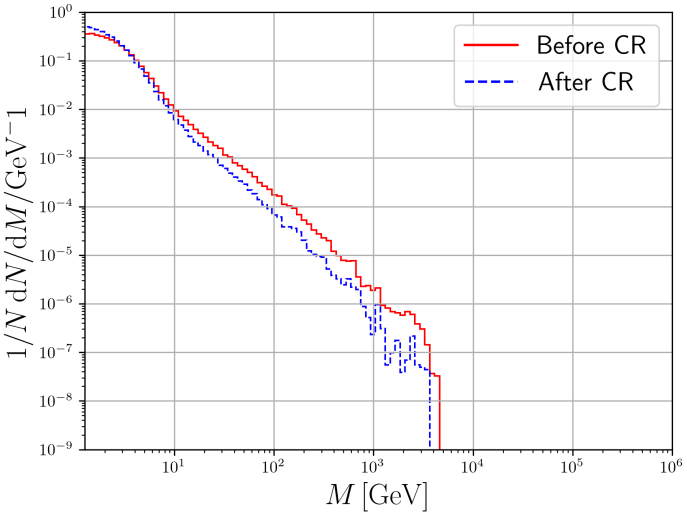

Figure 3.6: Distribution of invariant cluster masses before and after colour reconnection for MB pp collisions with centre-of-mass energy of $\sqrt{s} = 7\,\mathrm{TeV}$.

The statistical colour reconnection model, on the other hand, uses a simulated annealing algorithm [73] to find the configuration of clusters that results in the absolute lowest value of the colour length λ. It was found in Ref. [72] that the statistical colour reconnection prefers a quick cooling that does not result in a global minimum of λ in order to describe the data. In a recent paper the colour reconnection model was further modified such that the reconnections resulting in a colour-singlet gluon after hadronization are forbidden [74].

With increasing precision from the LHC and measurements becoming more and more differential, it becomes apparent that many non-perturbative aspects of particle production are far from being understood. In particular the description of the transition from the deconfined state to final state particles observed in the detectors remains difficult and only approachable by means of non-perturbative models. In this regard, Monte Carlo event generators provide an essential tool to study and evaluate different non-perturbative models. Especially the description of charged particle distributions in the high-multiplicity region [67], observables related to the distribution of hadrons containing strange quarks, such as Kaons and Λ-baryons [75], and the description of collective effects in *pp* collisions as observed in Ref. [76], seem to be out of reach for most Monte Carlo event generators. To address these issues, new and improved models for hadronization and colour reconnection are necessary [77]. Although Herwig 7 is able to describe general properties of MB and UE data, the description of the mentioned aspects is so far unsatisfactory. To account for the description of such observables we see the need to introduce a new model for colour reconnection and furthermore modify the strangeness production mechanism during hadronization.

CHAPTER 4

Baryonic Colour Reconnection

The significance of non-perturbative models has become apparent with a striking observation, made by the ALICE collaboration, which showed signs of enhancement of (multi-)strange hadrons in high-multiplicity pp events [76]. A property formerly only attributed to pA and AA collisions [78–83]. Possible explanations of these effects are rooted in the prospect that partonic matter shows some collective behaviour as in a hydrodynamical description, see e.g. Ref. [84]. Another possibility to account for strong and possibly long-range correlations between different hard partons in a single interaction can be introduced via colour reconnections [85]. One approach in this regard was followed e.g. in the Dipsy rope model [86, 87] where many overlapping strings are combined into a colour field of a higher representation which, in the Lund string model language, leads to an increased probability of producing strangeness during string fragmentation. In Pythia, this issue has been addressed with a modified fragmentation model, that allows the creation of string junctions resulting in a better description of data [77]. Further work in this regard focuses on a model for thermodynamical string fragmentation [88], where the main result is a shift of the transverse momentum of heavier particles. An accurate description of MB and UE observables within Herwig 7, has been achieved with the development of a new model for soft and diffractive interactions [53]. In this model, the importance of colour reconnections has already been observed for the description of high-multiplicity events. It was noted, that the mere production of baryons would lead to a reduction of charged multiplicity in favour of a rise of the multiplicity of heavier particles, effectively trading off mesons with baryons. For a review of the hadronization model and colour reconnection, we refer to Ch. 3. In this chapter, we propose an extension to the model for colour reconnection to account for the production of baryons to which we refer as *baryonic* colour reconnection. At the same time we reconsider the strangeness production mechanism and introduce an additional source of strangeness during the non-perturbative gluon splitting stage. With this modification to the hadronization model, we find that we can significantly improve the production rates of strange mesons as well as strange baryons. We compare the effects of these changes to observations made by various collider experiments, while the

description of general properties of the event, like charged-particle multiplicities and ratios of identified hadrons are of main interest, we furthermore compare our model to the ALICE strangeness-enhancement data.

This chapter is organized as follows: in Sec. 4.1 we review the physics underlying the model for baryonic colour reconnection. In Sec. 4.2 we explain the implementation of the algorithm and show with a simple proof of concept that the model indeed has the desired properties. In Sec. 4.3 we tune the free parameters of the model to various MB data and in Sec. 4.4 we make an extensive comparison to several important MB observables before highlighting the need for an additional strangeness production mechanism on the level of gluon splitting. In Sec. 4.5 we summarize our results.

4.1 The Baryonic Colour Reconnection Model

The only constraint upon forming a cluster is that the cluster has to be able to form a colourless singlet under $SU(3)$. In $SU(3)$ a coloured quark is represented as a triplet (3) and an anticoloured antiquark is represented as an anti-triplet ($\bar{3}$). Two triplets can be represented as an anti-triplet, and two anti-triplets can be represented as a triplet,

$$3 \otimes 3 = 6 \oplus \bar{3}, \tag{4.1}$$
$$\bar{3} \otimes \bar{3} = \bar{6} \oplus 3. \tag{4.2}$$

The clusters are a combination of these coloured quarks were only combinations are allowed that result in a colourless singlet. Here we consider the following allowed cluster configurations based on the $SU(3)$ structure of QCD. We begin with the normal cluster configuration which will be referred to as a *mesonic* cluster

$$3 \otimes \bar{3} = 8 \oplus 1. \tag{4.3}$$

In strict $SU(3)$ the probability of two quarks having the correct colours to form a singlet would be 1/9. Next we consider possible extensions to the colour reconnection that allows us to form clusters made out of 3 quarks. A *baryonic* cluster consists of three quarks or three antiquarks where the possible representations are,

$$3 \otimes 3 \otimes 3 = 10 \oplus 8 \oplus 8 \oplus 1, \tag{4.4}$$
$$\bar{3} \otimes \bar{3} \otimes \bar{3} = 10 \oplus 8 \oplus 8 \oplus 1. \tag{4.5}$$

In full $SU(3)$ the probability to form a singlet made out of three quarks would be 1/27. In the following we will introduce the algorithm we used for the alternative colour reconnection model. In order to extend the current colour reconnection model, which only deals with *mesonic* clusters, we allow the reconnection algorithm to find configurations that would result in a *baryonic* cluster.

4.2 Algorithm

As explained in Sec. 3.7, the colour reconnection algorithms in Herwig 7 are implemented in such a way that they lower the sum of invariant cluster masses. For clusters containing 3 constituents, the larger invariant cluster mass makes such a condition no longer sensible. As an alternative to the minimization of cluster masses, we consider a

simple geometric picture of nearest neighbours were we try to find quarks that approximately populate the same phase space region based on their rapidity y. The rapidity y is defined as

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right), \tag{4.6}$$

and is usually calculated with respect to the z-axis. As with the other colour reconnection algorithms, the starting point for the new rapidity based algorithm is the predefined colour configuration that emerges once the perturbative evolution of the parton shower has terminated and the remaining gluons are split non-perturbatively into quark-antiquark pairs, after which a list of clusters is created from all colour connected quarks and antiquarks. The final algorithm to form clusters of a baryonic type then consists of the following steps:

1. Shuffle the list of clusters to prevent the bias that comes from the order in which we consider the clusters for reconnection.

2. Pick a cluster (A) from that list and boost into the rest-frame of that cluster. The two constituents of the cluster $(q_\mathrm{A}, \bar{q}_\mathrm{A})$ are now in a back-to-back configuration. We define the direction of the antiquark as the positive z-direction of the quark axis.

3. Perform a loop over all remaining clusters and calculate the rapidity of the cluster constituents with respect to the quark axis in the rest frame of the original cluster for each other cluster in that list (B).

4. Depending on the rapidities the constituents of the cluster $(q_\mathrm{B}, \bar{q}_\mathrm{B})$ fall into one of three categories:

   - Mesonic: $y(q_\mathrm{B}) > 0 > y(\bar{q}_\mathrm{B})$,
   - Baryonic: $y(\bar{q}_\mathrm{B}) > 0 > y(q_\mathrm{B})$,
   - Neither.

   If the cluster neither falls into the mesonic, nor in the baryonic category listed above the cluster is not considered for reconnection.

5. The category and the absolute value, $|y(q_\mathrm{B})| + |y(\bar{q}_\mathrm{B})|$, for the clusters with the two largest sums are saved (these are clusters B and C in the following).

6. Consider the clusters for reconnection depending on their category. If the two clusters with the largest sum (B and C) are in the category *baryonic*, consider them for baryonic reconnection (together with cluster A) with probability p_B. If the category of the cluster with the largest sum is *mesonic* then consider it for normal reconnection with probability p_R. If a baryonic reconnection occurs, remove these clusters (A, B, C) from the list and do not consider them for further reconnections. A picture of the rapidity based reconnection for a mesonic configuration is shown in Fig. 4.1 and a simplified sketch for baryonic reconnection is shown in Fig. 4.2.

7. Repeat steps with the next cluster in the list.

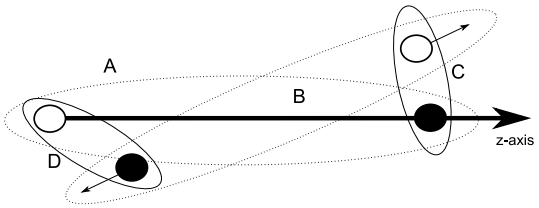

Figure 4.1: Representation of rapidity based colour reconnection where the quark axis of one cluster is defined as the z-axis in respect to which the rapidities of the constituents from the possible reconnection candidate are calculated. (A) and (B) are the original clusters. (C) and (D) would be the new clusters after the reconnection. Taken from Ref. [7].

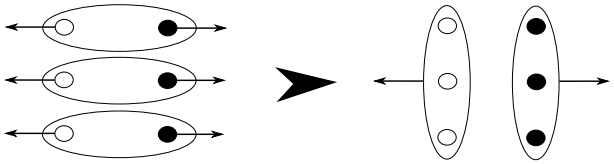

Figure 4.2: Configuration of clusters that might lead to baryonic reconnection. The small black arrows indicate the direction of the quarks. A reconnection is considered if all quarks move in the same direction and all antiquarks move in the same direction. Taken from Ref. [7].

We note that with this description we potentially exclude clusters from reconnection where both constituents have a configuration like $y(q_{\mathrm{B}}) > y(\bar{q}_{\mathrm{B}}) > 0$ w.r.t. the quark axis but assume that these clusters already contain constituents which are close in rapidity. The exclusion of baryonically reconnected clusters from further re-reconnection biases the algorithm towards the creation of baryonic clusters whose constituents are not the overall nearest neighbours in rapidity.

The extension to the colour reconnection model gives Herwig 7 an additional possibility to produce baryons on a different, more elementary level than on the level of cluster fission and cluster decay. In pp collisions with enhanced activity from multiple partonic interactions, a high density of clusters leads to an increased probability of finding clusters that are suitable for baryonic colour reconnection. We expect this model therefore to have a significant effect on charged-hadron multiplicities, especially on the high-multiplicity region. We also expect the model to have a significant impact on baryon and meson production since the model for baryonic colour reconnection effectively increases the multiplicity of baryons and decreases the multiplicity of mesons. In the following we show with a simple proof of concept by running Herwig 7 in combination with the baryonic colour reconnection model for different values of the baryonic reconnection probability, ($p_{\mathrm{B}} = 0.2, 0.5$), if our expectations do indeed get met by the new model. In Figs. 4.3 and 4.4 we see the influence of the new baryonic colour reconnection model for the different values of p_{B} on the charged-particle multiplicities and the $p_{\perp}$ spectra of $\pi^{+}+\pi^{-}$ and $\mathrm{p}+\bar{\mathrm{p}}$ yields in inelastic pp collisions at $\sqrt{s} = 7\,\mathrm{TeV}$ in the central rapidity region. As expected the model influences the hadronic multiplicities for large N_{ch} significantly. A larger baryonic reconnection probability reduces the number of high multiplicity events and shifts them towards lower multiplicities. The $p_{\perp}$ distribution of the $\pi^{+}+\pi^{-}$ shows an overall reduction while the $p_{\perp}$ spectra of the $\mathrm{p}+\bar{\mathrm{p}}$ shows

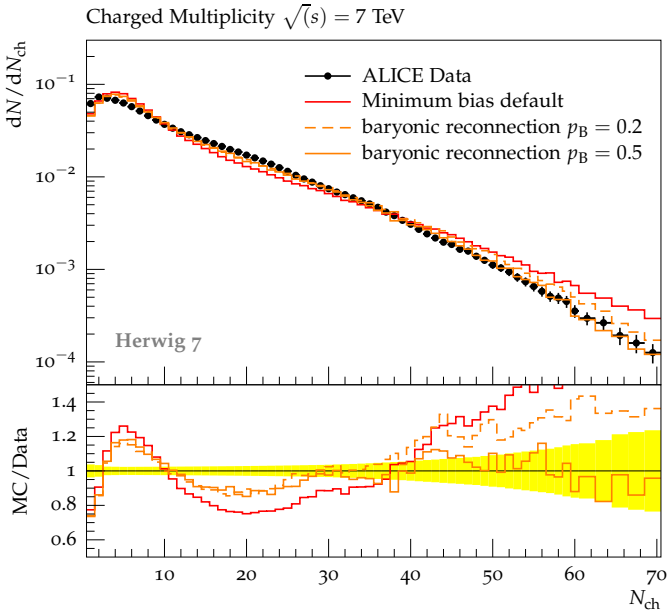

Figure 4.3: Measurement of the charged-particle multiplicity at $\sqrt{s} = 7\,\text{TeV}$ with ALICE at LHC [89]. Shown is a comparison of the new colour reconnection model for different reconnection probabilities with the default model of Herwig 7.

an overall enhancement due to baryonic colour reconnection. While the description of the low $p_\perp$ region improves, there are too many $\text{p} + \bar{\text{p}}$ with a $p_\perp > 2.5\,\text{GeV}$.

4.3 Tuning

The tuning is achieved by using the Rivet and Professor framework for Monte-Carlo event generators [91, 92]. In a first tuning attempt we keep the hadronization parameters that were tuned to LEP data at their default values and follow a similar tuning procedure as in Ref. [53]. We re-tune the main parameters of the model for multiple parton interactions in Herwig 7, the $p_{\perp,0}^{\text{min}}$ parameter and the inverse proton radius squared μ^2 in combination with the two parameters for colour reconnection p_R and p_B. In order to capture general features of MB observables we tune the model to a large variety of MB data from the ATLAS and ALICE collaborations at $\sqrt{s} = 7\,\text{TeV}$ [67,90]. The following observables were used with equal weights:

- the pseudorapidity distributions for $N_\text{ch} \geq 1$, $N_\text{ch} \geq 2$, $N_\text{ch} \geq 6$, $N_\text{ch} \geq 20$,
- the transverse momentum of charged particles for $N_\text{ch} \geq 1$,
- the charged particle multiplicity for $N_\text{ch} \geq 2$,
- the mean charged transverse momentum vs. the multiplicity of charged particles for $p_\perp > 500\,\text{MeV}$ and $p_\perp > 100\,\text{MeV}$,
- the pion and the proton yield in the central rapidity region $|y| < 0.5$.

The outcome of this tune is listed in Tab. 4.1 where we show the parameter values that resulted in the lowest value of χ^2/N_dof and the values from the default tune of Herwig 7.1 without the baryonic colour reconnection model. The change in the colour reconnection

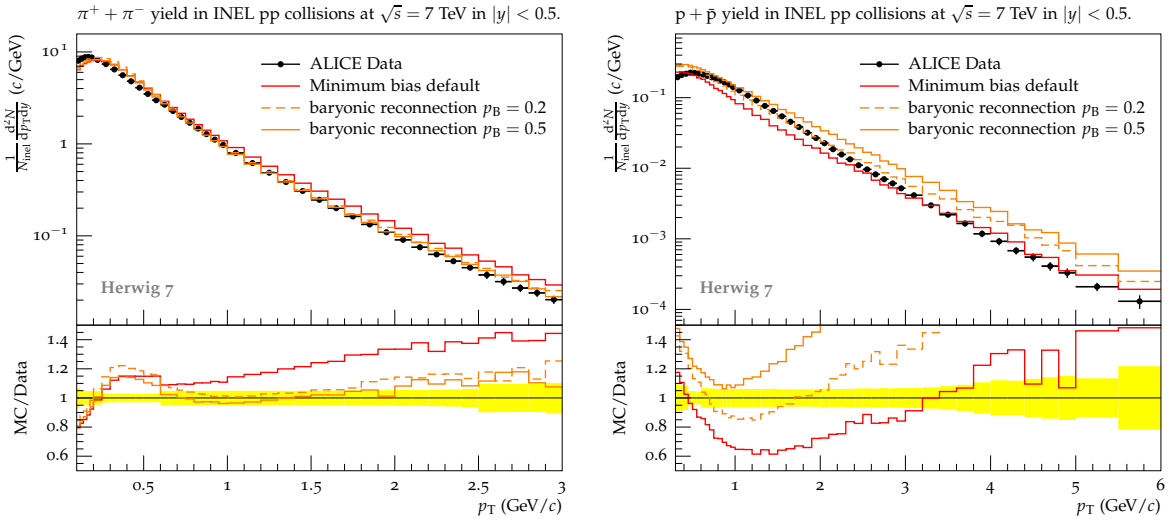

Figure 4.4: The transverse momentum spectra for $\pi^+ + \pi^-$ and $p + \bar{p}$ as measured by ALICE at $\sqrt{s} = 7\,\text{TeV}$ [90] in the central rapidity region $|y| < 0.5$.

algorithm and the possibility to produce baryonic clusters results in an overall better description of the considered observables. While still being able to accurately describe MB data we see the expected improvement in the charged multiplicity distributions for the high-multiplicity region. The results of the tuning will be presented and discussed in the next section.

	$p_{\perp,0}^{\min}/\text{GeV}$	$\mu^2/\,\text{GeV}^2$	p_R	p_B
default	3.502	1.402	0.5	0
tune	3.269	1.963	0.543	0.2086

Table 4.1: Results of the parameter values from the tuning procedure that resulted in the smallest χ^2/N_{dof} value for $\sqrt{s} = 7\,\text{TeV}$ centre-of-mass energy compared with the default tune from Herwig 7.1.

4.4 Results

Changes in the colour reconnection model are always deeply tied with the peculiarities of the hadronization model. In principle we would have to re-tune all parameters that govern the hadronization model in Herwig 7. This is usually done in a very dedicated and long study with data from LEP experiments [3, 93]. We propose a simplified procedure since little to no changes are expected with the extension to the colour reconnection model in the e^+e^- environment. At LEP, the colour structure of an event is not changed significantly through colour reconnection since it is already well defined by the parton shower. This was confirmed by comparing the new model to a wide range of experimental data from LEP. We therefore keep the hadronization parameters that were tuned to LEP data (see Refs. [3, 4]) at their default values. We also note that this does not replace a dedicated study concerned with the tuning and validation of hadronization parameters. Especially at pp collisions a different model

for colour reconnection leads to changes in the interplay between the clusters and the hadronization in an unforeseeable way. Nonetheless we restrain ourselves to the explained simplified method in order to make qualitative statements about the new model for colour reconnection. The new model with the tuned parameters improves the description of all observables considered in the tuning procedure. The effect of the baryonic colour reconnection was already demonstrated in Fig. 4.3. In Fig. 4.5 we show the same distribution of the charged-particle multiplicity for the central region $|y| < 1$ with the tuned parameter values. Again we see the expected fall off for high multiplicities. The baryonic colour reconnection model is able to describe the whole region fairly well compared to the plain colour reconnection model. Only the low multiplicity region $n < 10$ is overestimated by a factor of $\approx 10\%$ and for $n < 5$ underestimated. In Fig. 4.5 we also show a similar observable for a wider rapidity region $|y| < 2.4$ and up to $n = 200$ as measured by CMS [67]. Again the central multiplicity region shows a significant improvement. For multiplicities $n > 80$ we note a slight overestimation of the data but are still within error bars.

This can be understood quite simply: the more activity in an event, the more likely it becomes that a cluster configuration leading to a baryonic reconnection is found. The high-multiplicity events therefore exhibit a disproportionately large fraction of baryonic reconnections. Due to the highly restricted phase space for the production of baryons from baryonic clusters, fewer particles are produced than with mesonic clusters of the same invariant mass. This in return lowers the charged multiplicity.

4.4.1 Strangeness production

We also observe the proposed change in mesonic and baryonic activity in the $p_\perp$ spectra of pions and protons. Especially the p/π ratio and the $p_\perp$ distributions improve significantly which should be considered first in a model that tries to describe flavour multiplicities. When looking at the $p_\perp$ distributions of K and Λ we see that none of the performed tunes are able to capture the essence of these distributions correctly. This is no surprise since we have not altered the production mechanism of strangeness. We merely observe a small increase in the $p_\perp$ distribution of Λ baryons.

In the non-perturbative regime of Herwig 7, strange quarks can be produced at the stage of cluster fissioning and cluster decay, controled by the parameter `PwtSquark`. The production of strange quarks during cluster decay additionally depends on the available phase space (see Sec. 3.3). In order to account for an additional source of strangeness in the event generator work flow, we allow the gluons which remain at the end of the parton shower evolution to split non-perturbatively into strange quark-antiquark pairs. To allow the production of strangeness during gluon splitting, the non-perturbative gluon mass has to be adjusted such that $m_g > 2m_s$. This has shown to have no significant effects on the description of LEP data. To control the production weight of strangeness during gluon splitting we introduce the parameter `SplitPwtSquark`. In a second tuning attempt we additionally tune the parameters `PwtSquark` and `SplitPwtSquark`, which are resposible for the strangeness production during hadronization, to the $p_\perp$ distribution of the $\pi^+\pi^-$, $\mathrm{K}^+ + \mathrm{K}^-$, $\mathrm{p} + \bar{\mathrm{p}}$ yields in inelastic pp collisions at 7 TeV [90] and the $p_\perp$ distribution of Λ [75]. The resulting parameter values are listed in Tab. 4.2. We see a significant improvement for many hadron flavour observables which were not considered in the tuning. In order to study the different effects from the baryonic colour

	$p_{\perp,0}^{\min}/\mathrm{GeV}$	$\mu^2/\,\mathrm{GeV}^2$	p_R	p_B	PwtSquark	SplitPwtSquark
default	3.502	1.402	0.5	0	0.665	0
tune	3.053	1.282	0.772	0.477	0.291	0.824

Table 4.2: Results of the parameter values from the tuning procedure that resulted in the smallest χ^2/N_dof value for $\sqrt{s} = 7\,\mathrm{TeV}$ compared with the default tune from Herwig 7.1.

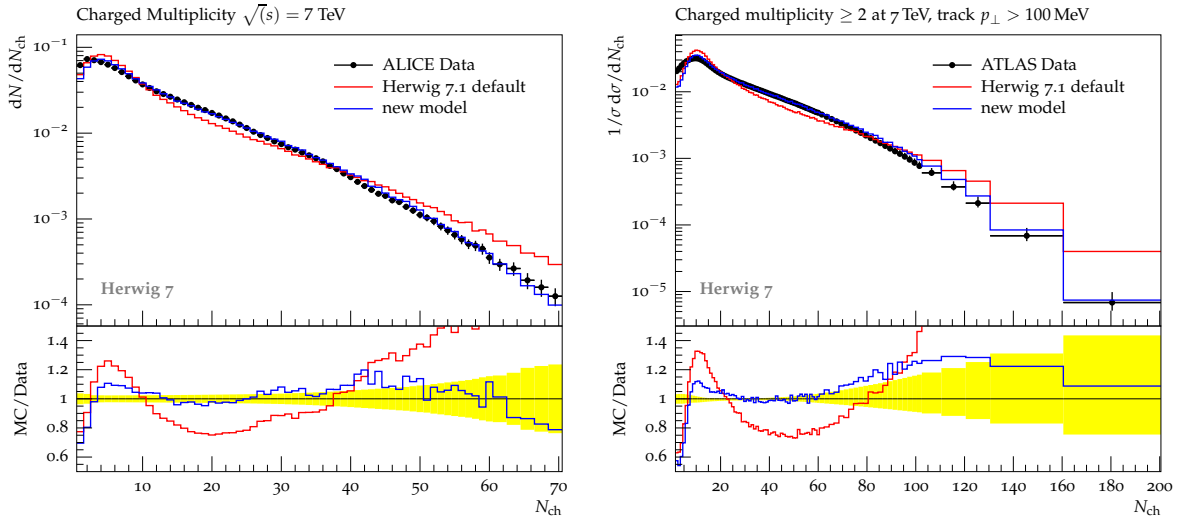

Figure 4.5: Multiplicity distributions as measured by ALICE for the central rapidity region $|\eta| < 1$ up to $N_{\mathrm{ch}=70}$ [89] and ATLAS for $|\eta| < 2.4$ up to $N_\mathrm{ch} = 200$ for all particles with $p_\perp > 100\,\mathrm{MeV}$ [67]. We compare the plain colour reconnection model (default) with the new baryonic colour reconnection model.

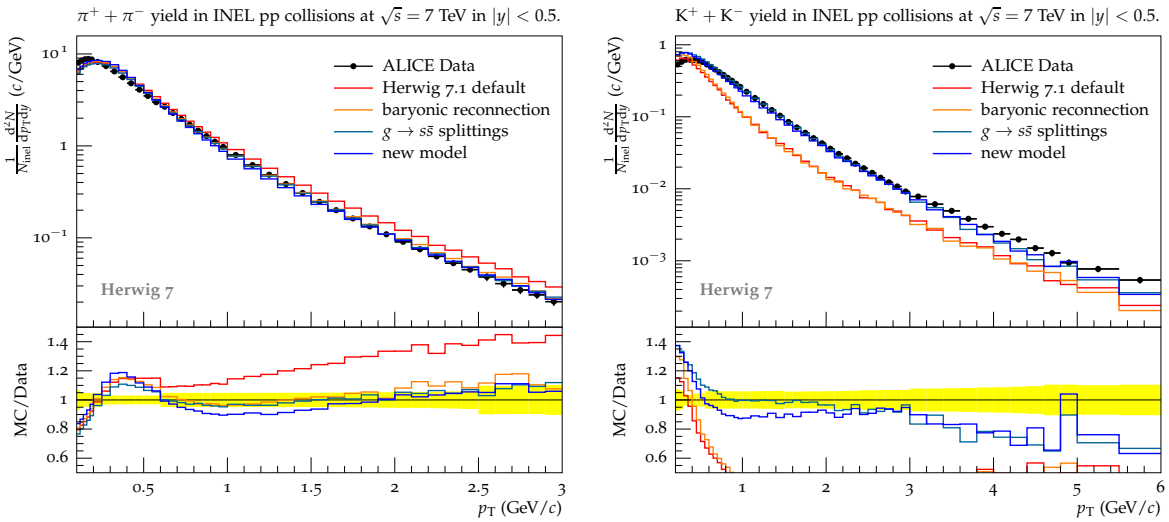

Figure 4.6: The transverse momentum spectra for $\pi^+ + \pi^-$ and $K^+ + K^-$ as measured by ALICE at $\sqrt{s} = 7$ TeV [90] in the central rapidity region $|y| < 0.5$.

reconnection model in combination with the possibility to produce strange quarks during gluon splitting we compare runs with the default model (Herwig 7.1 default), the baryonic colour reconnection model (baryonic reconnection), one run where we allow the gluons to split into strange quarks ($g \to s\bar{s}$ splittings) in combination with the plain colour reconnection model and a run where we combine the model for baryonic colour reconnection with the modified gluon splitting with the tuned parameters (new model). In Fig. 4.6 we show the $p_\perp$ distributions of π and K in the central rapidity region as measured by ALICE [90] and in Fig. 4.7 the corresponding $p + \bar{p}$ distribution. While all options improve the description of pions we see that the K distribution can only be described if we take the additional source of strangeness into account. The proton $p_\perp$ distribution is mainly driven by baryonic reconnection. The rate increases for all $p_\perp$ regions but we overestimate the data by a large factor for $p_\perp > 3\,\text{GeV}$ and for the very low $p_\perp$ region. Since all options show the same trend this might indicate some problems with the hard part of the model for multiple parton interactions which dominates the production of particles for higher values of $p_\perp$. In Fig. 4.8 we consider the hadron ratios K/π, and p/π. The new model does a significantly better job in describing the data and only the combined effect of the enhanced baryon production through the change in the colour reconnection model and gluon splitting into strange quarks is able to give a satisfying description of both observables.

In Figs. 4.9, 4.10, 4.11 and 4.12 we compare the model to $\sqrt{s} = 7$ TeV data from CMS [75] for the strange flavour observables of K^0_S, Λ and Ξ^-. The new model improves the description for all observables published in this analysis. Again we show the effects of the different contributions and note that the best description can only be achieved with a combination of baryonic colour reconnection and gluon splitting into strange quarks (new-tune). The Λ/K^0_S distribution shows a good description in the turn on region but the high $p_\perp$ tail is not well described. A similar observation was made with Pythia in Ref. [77]. Surprisingly the Ξ^-/Λ distribution is able to capture the general trend but due to large errors in the high $p_\perp$ region no sound conclusion can be drawn. Especially the rapidity distributions and the ratios Λ/K^0_S and Ξ^-/Λ show an improvement compared to the default model. Again we point out the interplay

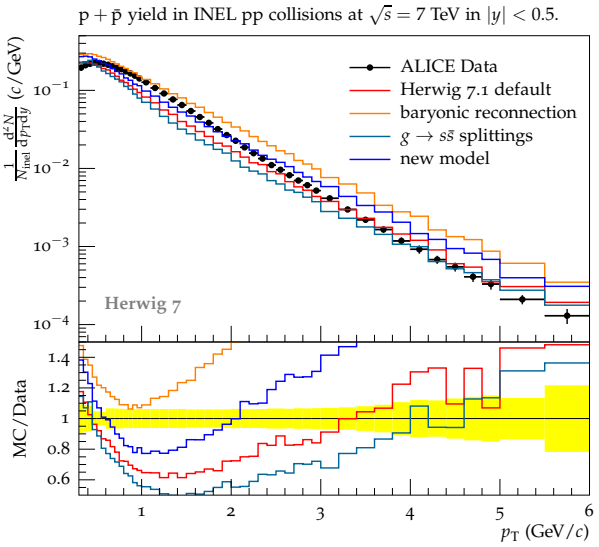

Figure 4.7: The transverse momentum spectrum for $\mathrm{p} + \bar{\mathrm{p}}$ as measured by ALICE at $\sqrt{s} = 7$ TeV [90] in the central rapidity region $|y| < 0.5$.

between baryonic colour reconnection and the strangeness production mechanism which is responsible for the improvement in the description of the heavy baryons Λ and Ξ^-.

In an analysis by ALICE [76], an enhancement of strange to non-strange hadron production with increasing particle multiplicity in *pp* collisions was observed. It was noted that the observed enhancement increases with the strangeness content of the hadron and not with the mass or baryon number. This presented a novelty in *pp* collisions since these effects were formerly only attributed to the high density systems of *pPb* and *PbPb* collisions. Since we are developing a model that incorporates strangeness production and the enhanced production of baryons it is instructive to compare our model to the data published by the ALICE collaboration. In Figs. 4.13, 4.14 we show the ratio of the yields of K^0_S, Λ, Ξ and Ω to the pion $(\pi^+ + \pi^-)$ yield as a function of $\langle \mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta \rangle$ at $\sqrt{s} = 7\,\mathrm{TeV}$. The new model describes the data significantly better than the default version of Herwig 7.1. The model manages to describe the rising trend of the strange to non-strange hadron production qualitatively for all observables. Only for $2K^0_S/(\pi^+ + \pi^-)$ we see a deviation for $\langle \mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta \rangle < 4$. In Fig. 4.15 we show the ratio of the single strange hadrons $(\Lambda + \bar{\Lambda})/2\mathrm{K}^0_{\mathrm{s}}$ and the ratio of non strange hadrons $(\mathrm{p} + \bar{\mathrm{p}})/(\pi^+ + \pi^-)$. These yield ratios do not change significantly with multiplicity. The new model for baryonic colour reconnection in combination with the $g \to s\bar{s}$ splitting mechanism, describes the data significantly better than the default version of Herwig 7.1.

We highlight that the new model was tuned to general aspects of MB data and not to the data from the ALICE study. The new model manages to capture the observed trend in all observables considered and describes the data qualitatively to a good accuracy. The model presents an additional possibility to understand and describe the enhanced production of strangeness in high-multiplicity events from a microscopical point of view.

4.4.2 Spectra of cluster masses

In this section we discuss the effects of the baryonic colour reconnection model on the clusters and the distribution of cluster masses. This is done for non-diffractive

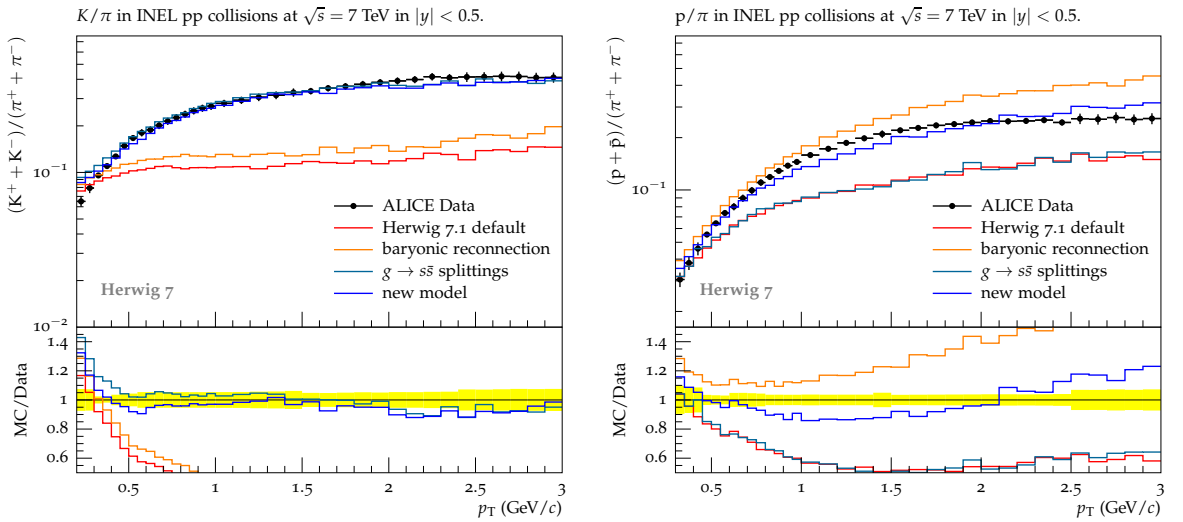

Figure 4.8: Transverse momentum spectra for the ratios p/π and K/π as measured by ALICE at $\sqrt{s} = 7$ TeV in the very central rapidity region $|y| < 0.5$ [90].

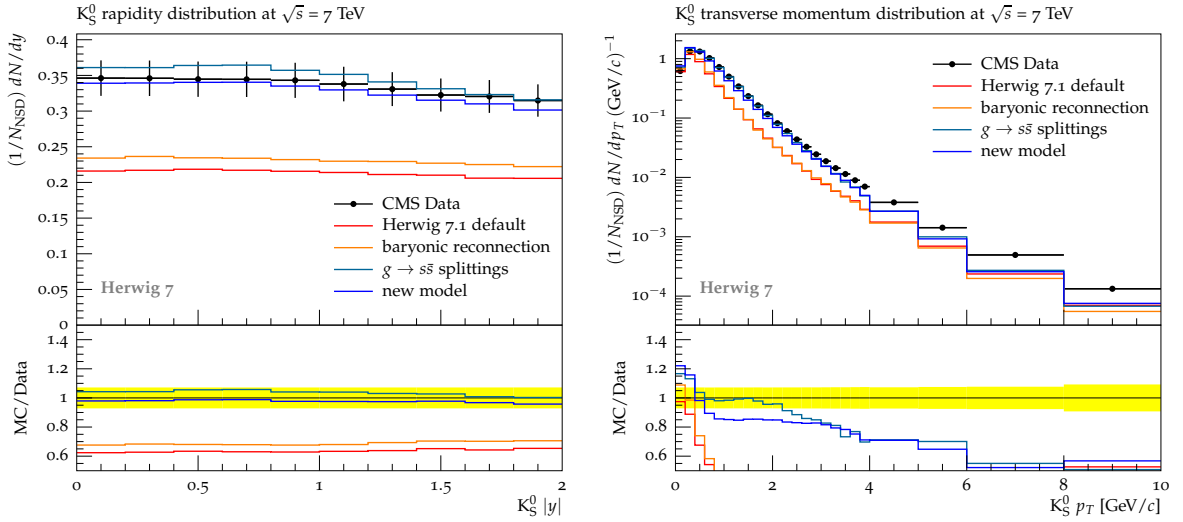

Figure 4.9: The K^0_S rapidity and $p_\perp$ distribution as measured by CMS at $\sqrt{s} = 7$ TeV [75].

processes, since colour reconnection has no effect on the simulation of diffraction. In Fig. 4.16 we show the effects of the baryonic colour reconnection model on the distribution of cluster masses. It can be seen that after colour reconnection the cluster masses get shifted towards smaller values, although the effect is not as severe as in the plain colour reconnection model (see Fig. 4.17). In a direct comparison between the plain colour reconnection model and the baryonic colour reconnection model we see that the baryonic colour reconnection model favours the production of heavier clusters. In Fig. 4.18 we show the distribution of cluster masses after colour reconnection separately for baryonic and mesonic clusters. The contribution in the high-mass region mainly comes from mesonic clusters. Baryonic clusters dominate the mid-mass region between 1 and 13 GeV while large baryonic clusters are highly suppressed. With this in mind, a picture of the cluster configuration emerges which, in order to be able to describe the data, favours the production of baryonic clusters with an intermediate cluster mass and small fluctuations towards clusters of very high masses. In general, one can say

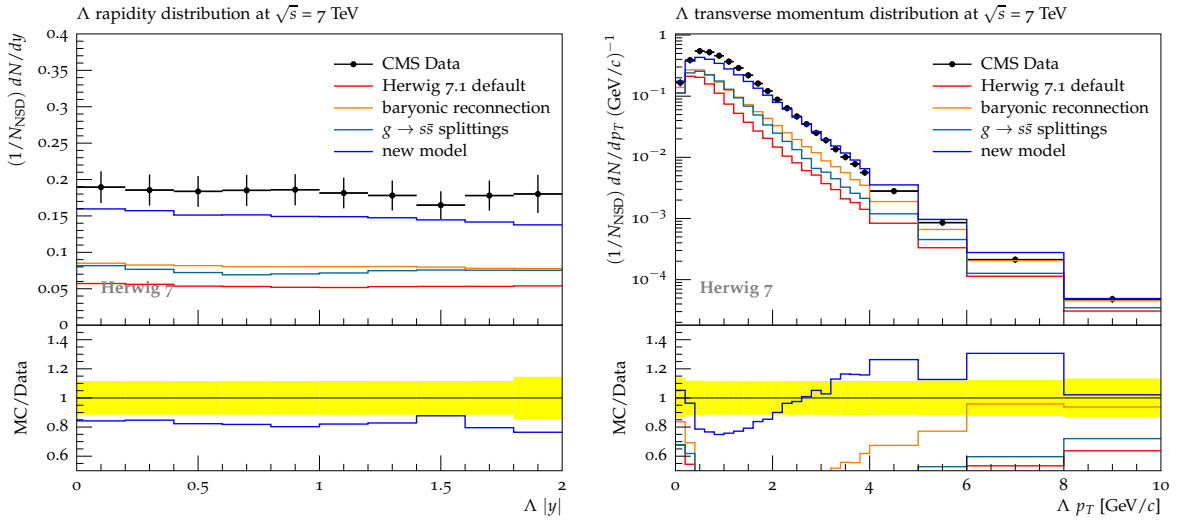

Figure 4.10: The Λ rapidity and $p_\perp$ distribution as measured by CMS at $\sqrt{s}$ = 7 TeV [75].

that lighter clusters lead to fewer heavy particles due to the highly restricted phase space during cluster decay. This makes the plain model for colour reconnection, which is based on the direct reduction of cluster masses not able to reproduce the observables concerned with heavier particles due to a lack of heavy clusters.

4.5 Conclusion

We have implemented a new model for colour reconnection which is entirely based on a geometrical picture instead of an algorithm that tries to directly minimize the invariant cluster mass. Additionally we allow reconnections between multiple mesonic clusters to form baryonic clusters which was not possible in the plain colour reconnection model. This mechanism results in a lever on the baryon to meson ratio which is a necessary starting point to describe flavour observables. The amount of reconnection also depends on the multiplicity of the event. To account for strangeness production we allow for non-perturbative gluon splitting into strange quark-antiquark pairs. Only with this additional source of strangeness, it is possible to get a reasonably good description of the $p_\perp$ spectra of the kaons. The description of heavy baryons Λ, and Ξ^-, improves once we combine the new baryonic colour reconnection model and the additional source of strangeness. The model was tuned to 7 TeV MB data and various hadron flavour observables. The full range of MB data can be described with a similar good quality as the default model of Herwig 7.1. Additionally improve the description of hadron flavour observables significantly.

A comparison with ALICE data concerning the enhancement of (multi-)strange hadrons led us to the conclusion that our simple model is able to reproduce the trend of the observables qualitatively and in some regions with high accuracy. We have shown that the relative increase in strange hadron fraction with respect to multiplicity can also be described by a microscopic approach introducing correlations in the final state via colour reconnection in combination with the cluster hadronization model.

A shortcoming of the model lies in the algorithm which is biased by the order of clusters considered for reconnection and the fact that baryonic clusters cannot be re-

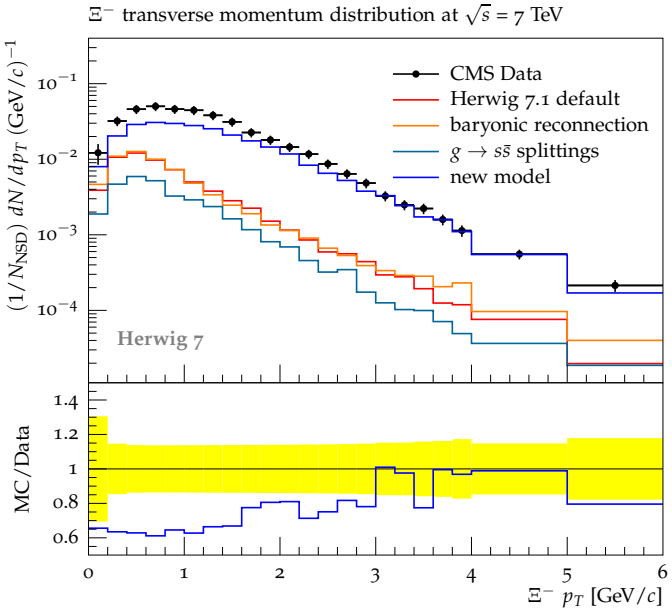

Figure 4.11: The Ξ^- $p_\perp$ distribution as measured by CMS at $\sqrt{s} = 7$ TeV [75].

reconnected. This will ultimately yield clusters which do not consist of the nearest neighbours in phase space but a small overlap between the clusters will still be present. Due to the sheer amount of possibilities on how to assign baryonic clusters we are forced to introduce some sort of arbitrariness when it comes to the cluster assignment. When comparing the baryonic colour reconnection model with the plain colour reconnection model, we see that the baryonic colour reconnection model does not have the same effect on the invariant mass distribution in terms of reduction of cluster masses but fuses mesonic clusters together in order to form baryonic clusters and therefore adds an additional possibility to produce heavy baryons. According to the data a significant reduction in cluster mass is not favoured. The data prefers more fluctuations in cluster size and explicitly welcomes the possibility to produce baryonic clusters. Otherwise the production of heavy strange baryons is highly suppressed. Overall, we have shown that changes in the model for colour reconnection and the gluon-splitting mechanism have significant effects on the description of MB flavour observables.

The model for baryonic colour reconnection in combination with the modified gluon splitting was released with Herwig version 7.1.4, and became the default model with Herwig version 7.2 [94]

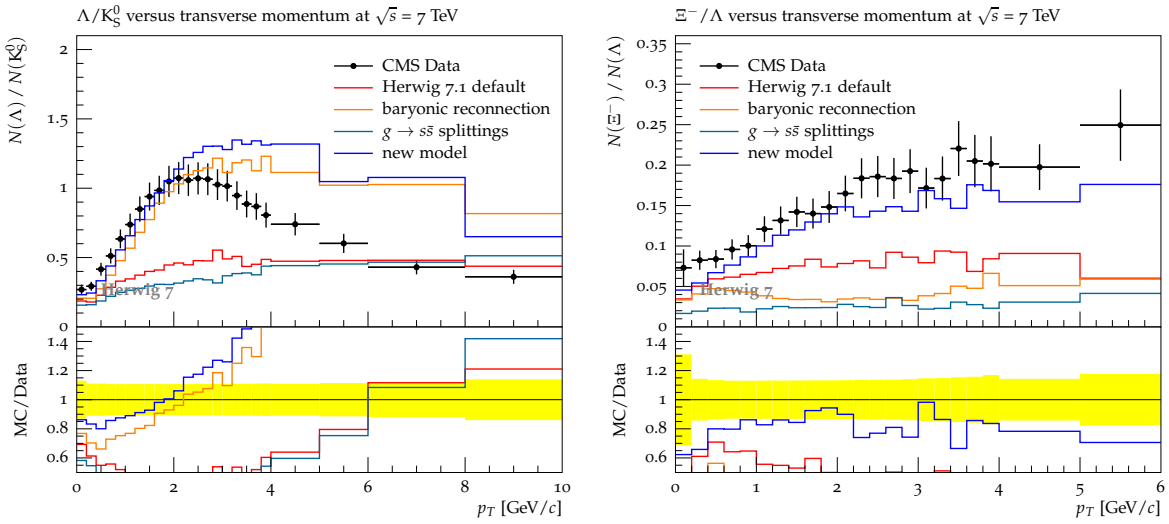

Figure 4.12: The Λ/K^0_S and the Ξ^-/Λ $p_\perp$ distribution as measured by CMS at $\sqrt{s} = 7$ TeV [75].

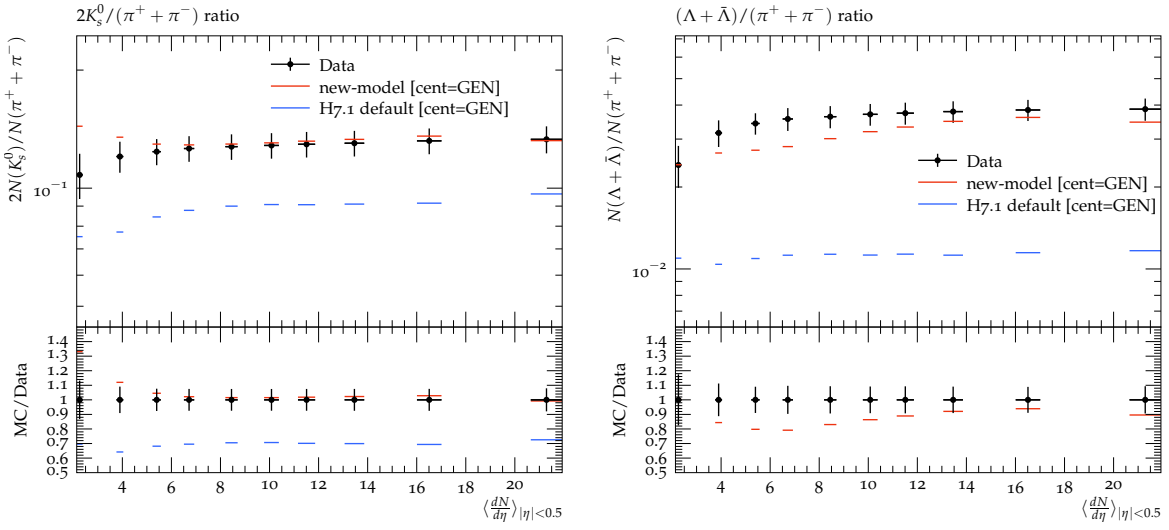

Figure 4.13: Integrated (multi)-strange particle yield ratios to $\pi^+ + \pi^-$ as a function of $\langle dN_{ch}/d\eta \rangle$ for $|\eta| < 0.5$. The values from the ALICE analysis [76] are compared to the default model of Herwig 7.1 and the new model.

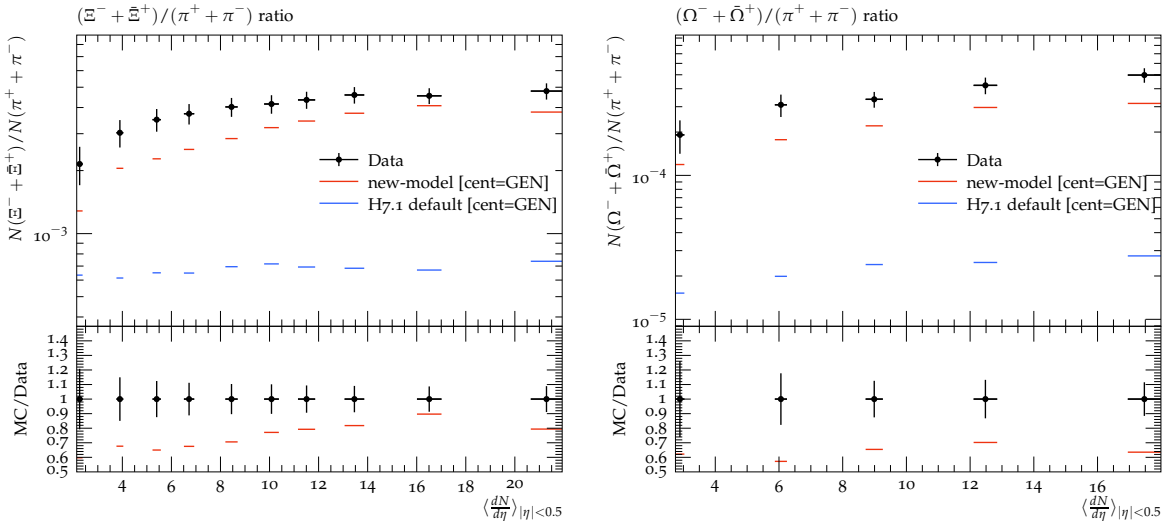

Figure 4.14: Integrated (multi)-strange particle yield ratios to $\pi^+ + \pi^-$ as a function of $\langle \mathrm{d}N_{ch}/\mathrm{d}\eta \rangle$ for $|\eta| < 0.5$. The values from the ALICE analysis [76] are compared to the default model of Herwig 7.1 and the new model.

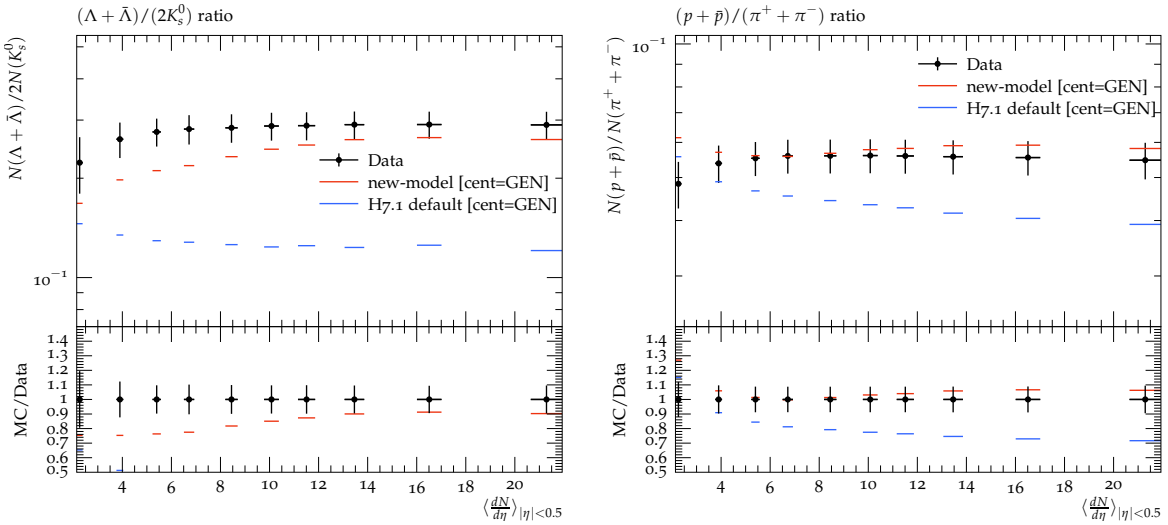

Figure 4.15: Integrated particle yield ratio of $(\Lambda + \bar{\Lambda})/2\mathrm{K}^0_\mathrm{s}$ and $(\mathrm{p} + \bar{\mathrm{p}})/(\pi^+ + \pi^-)$ as a function of $\langle \mathrm{d}N_{ch}/\mathrm{d}\eta \rangle$ for $|\eta| < 0.5$. The values from the ALICE analysis [76] are compared to the default model of Herwig 7.1 and the new model.

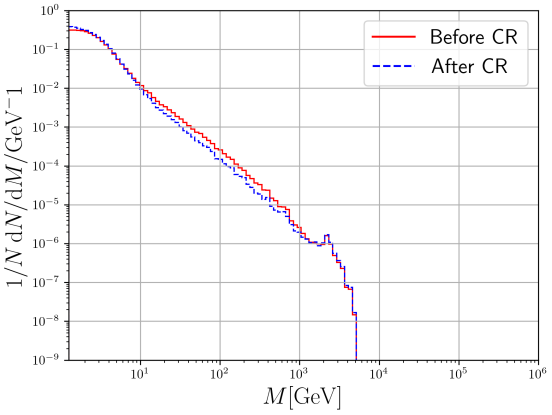

Figure 4.16: Comparison between the distributions of invariant cluster masses before and after the colour reconnection with the baryonic colour reconnection model.

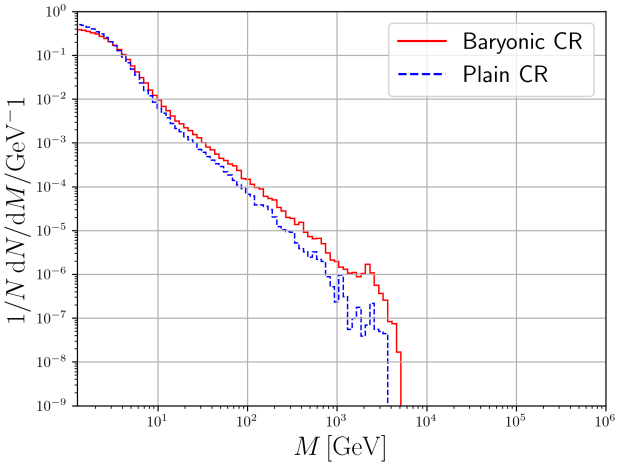

Figure 4.17: Comparison between the distribution of invariant cluster masses after colour reconnection for the plain colour reconnection model and the baryonic colour reconnection model.

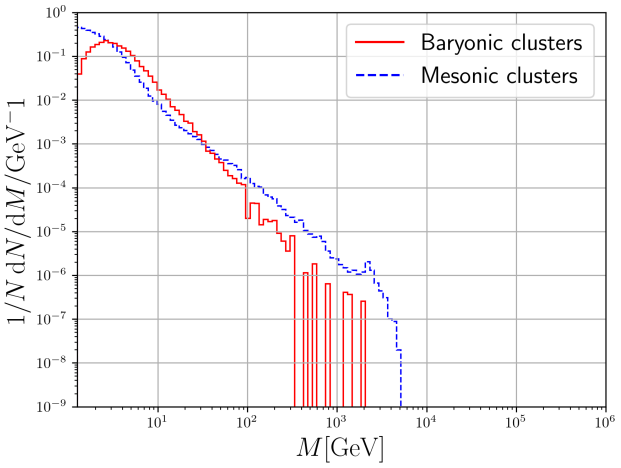

Figure 4.18: The invariant mass distribution of mesonic and baryonic clusters after colour reconnection.

CHAPTER 5

Colour Reconnection from Soft Gluon Evolution

Colour reconnection models have become a crucial part in phenomenological studies and are necessary for the description of MB and UE data at hadron colliders [72, 74, 87, 95–97]. The main goal of these models is to incorporate sub-leading colour effects on the level of hadronization. While the bulk of observables can be described with simple colour reconnection models, they quickly reach their limit of possibilities when confronted with more differential observables, for instance flavour composition with respect to event multiplicity [76]. The inability of most models to describe baryon production also led to improved models of colour reconnection [7, 77], one of which we introduced in Ch. 4.

While the spectrum of colour singlet cluster systems is typically predicted by the $N_c \to \infty$ parton shower evolution, the colour reconnection dynamics incorporates sub-leading colour effects which we expect to consist of a perturbative, as well as a non-perturbative component. While work is ongoing in analysing multi-parton emission dynamics beyond the leading colour approximation [98, 99], no consistent connection between these approaches to existing hadronization models has been made yet. In this chapter, we take the first step in this direction by approaching colour reconnection from a perturbative point of view and take the perturbative structure as a starting point to calculate possible colour reconnection effects at amplitude level. This results in a perturbatively inspired evolution of the colour-flow due to soft-gluon exchanges. We develop a toy Monte Carlo which allows us to study the colour evolution of a system of up to 5 clusters and furthermore analyze the properties of the resulting cluster configurations which are favoured by our perturbative ansatz and find strong support for geometric colour reconnection models.

This chapter is organized as follows. In Sec. 5.1 we review the details of colour preconfinement and the cluster hadronization model relevant for this chapter. In Sec. 5.2 we introduce the concept of perturbative colour evolution of a scattering amplitude which results in a mixing of possible colour structures in the (infrared) renormalized amplitude by means of a renormalization group equation in colour space. In Sec. 5.3 we

state a general algorithm which incorporates the perturbative colour-flow evolution in the form of a colour reconnection algorithm that calculates the probabilities to change from an initial colour-flow to a final colour-flow by calculating the overlaps of an evolved amplitude in colour space and a target colour structure. We also investigate the exactly solvable evolution of a two-cluster system in different kinematic regimes. In Sec. 5.4 we introduce the concept of a baryonic colour-flow and in Sec. 5.5 we present the numerical results for the full evolution of systems containing up to five clusters. In Sec. 5.6 we take the first step to isolate building blocks of our approach which might function as the basis of a new model for colour reconnection implementable in Monte Carlo event generators and in Sec. 5.7 we draw conclusions on our approach in terms of application and feasibility.

5.1 Pre-Confinement and Cluster Hadronization

The cluster hadronization model is an essential ingredient for Monte Carlo event generators such as Herwig 7 [3, 4] and Sherpa [2]. After the perturbative evolution of the parton shower has terminated, it converts partons from the scale of the parton shower infrared cutoff of order 1 GeV into the observed hadrons at energy scales of order Λ_{QCD}. The cluster hadronization model is based on the property of colour preconfinement which states that at any scale of the perturbative cutoff the colour structure of the parton shower is such that colour singlet combinations of partons can be formed with an asymptotically invariant mass distribution. The invariant mass distribution is independent of the properties of the hard scattering process or the parton shower itself.

After the evolution of the parton shower has terminated each coloured quark q is colour connected to an anticoloured antiquark $\bar{q}$ forming a colour singlet cluster. The properties of the clusters are solely determined by the individual quark flavours and the invariant cluster mass

$$M^2 = (p_q + p_{\bar{q}})^2 , \tag{5.1}$$

where $p_q, p_{\bar{q}}$ are the four-momenta of the constituent quarks. In the approach followed here, we neglect the specific flavours of the quarks since we are mainly interested in the mass distribution of the resulting clusters after colour evolution. Furthermore we stay in the massless parton limit, in which Eq. (5.1) simplifies to

$$M^2 = 2p_q \cdot p_{\bar{q}}. \tag{5.2}$$

5.2 Perturbative Colour Evolution

QCD scattering amplitudes are in general vectors in both spin and colour space. Since we are mainly interested in the colour structures we neglect the helicity degrees of freedom. A QCD scattering amplitude $\mathcal{M}$ can then be decomposed in a basis of contributing colour structures,

$$|\mathcal{M}\rangle = \sum_{\sigma} \mathcal{M}_\sigma |\sigma\rangle, \tag{5.3}$$

where σ represents the possible colour structures of the amplitude. We work in the colour-flow basis [100] in which colour structures can be labelled by permutations σ of the colour indices and describe how colour charge is flowing from one leg to another.

For n outgoing coloured legs and n outgoing anticoloured legs we use the notation from Ref. [99] to display which coloured and anticoloured legs are colour connected with each other

$$|\sigma\rangle = \left| \begin{array}{ccc} 1 & \cdots & n \\ \sigma(1) & \cdots & \sigma(n) \end{array} \right\rangle = \delta^{\alpha_1}_{\bar{\alpha}_{\sigma(1)}} \cdots \delta^{\alpha_n}_{\bar{\alpha}_{\sigma(n)}}, \tag{5.4}$$

where $\alpha_{1\ldots n}$ and $\bar{\alpha}_{\sigma(1)\ldots\sigma(n)}$ are fundamental and anti-fundamental indices assigned to the colour and anticolour legs respectively. They take values in the actual number of colours $1, ..., N_c$. Now we consider virtual corrections to the amplitude. Virtual corrections are in general both infrared (IR) and ultraviolet (UV) divergent. While UV divergences relate to the running of the strong coupling, IR divergences drive the evolution of a scattering amplitude. In this case the renormalization program can be used to sum large logarithmic contributions of infrared origin to all orders in perturbation theory [101]. In $d = 4 - 2\epsilon$ dimensions, the bare amplitude $|\tilde{\mathcal{M}}\rangle$ can be related to the renormalized amplitude $|\mathcal{M}\rangle$ in the following way

$$|\mathcal{M}(\{p\}, \mu^2)\rangle = \mathbf{Z}^{-1}(\{p\}, \mu^2, \epsilon)|\tilde{\mathcal{M}}(\{p\}, \epsilon)\rangle, \tag{5.5}$$

where $\{p\}$ is the set of outgoing momenta, $\epsilon = (d-4)/2$ is the dimensional regularization parameter in d dimensions, and μ^2 is the scale at which the IR renormalization has been performed. The renormalization constant $\mathbf{Z}$ is an operator in the space of colour structures and sums the IR divergences to all orders resulting in a finite renormalized amplitude.

By taking a logarithmic derivative of the bare amplitude with respect to the renormalization scale μ^2, we obtain the following evolution equation [101, 102]

$$\mu^2 \frac{\mathrm{d}}{\mathrm{d}\mu^2}|\mathcal{M}(\{p\}, \mu^2)\rangle = \mathbf{\Gamma}(\{p\}, \mu^2)|\mathcal{M}(\{p\}, \mu^2)\rangle, \tag{5.6}$$

where $\mathbf{\Gamma}(\{p\}, \mu^2)$ is the so-called soft anomalous dimension matrix

$$\mathbf{\Gamma}(\{p\}, \mu^2) = -\mathbf{Z}^{-1}(\{p\}, \mu^2, \epsilon)\mu^2 \frac{\partial}{\partial \mu^2}\mathbf{Z}(\{p\}, \mu^2, \epsilon). \tag{5.7}$$

This matrix encodes the residues of the $1/\epsilon$ divergences contained in $\mathbf{Z}$. Both $\mathbf{Z}$ and $\mathbf{\Gamma}$ are matrices in colour space and mix amplitudes with different colour structures. Solutions to the evolution equation take the general form

$$|\mathcal{M}(\{p\}, \mu^2)\rangle = \mathbf{U}(\{p\}, \mu^2, \{M^2_{ij}\})|\mathcal{H}(\{p\}, Q^2, \{M^2_{ij}\})\rangle, \tag{5.8}$$

where $\mathcal{H}(\{p\}, Q^2, \{M^2_{ij}\})$ represents the hard scattering amplitude before the evolution and

$$\mathbf{U}(\{p\}, \mu^2, \{M^2_{ij}) = \exp\left(-\int_{\mu^2}^{M^2_{ij}} \frac{\mathrm{d}q^2}{q^2}\mathbf{\Gamma}(\{p\}, \mu^2)\right), \tag{5.9}$$

is the evolution operator in the space of colour structures. Here we have neglected the ordered exponential since we assume that the soft anomalous dimension matrix commutes at different scales. We furthermore have not chosen a fixed scale to provide the initial conditions for the evolution but rather have chosen an upper limit on the

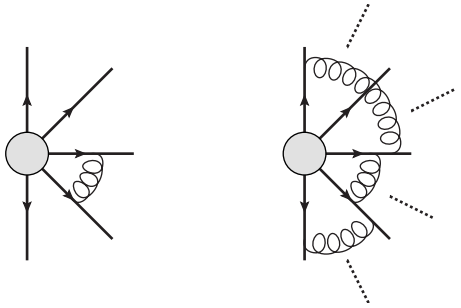

Figure 5.1: Diagrammatic representation of single soft gluon exchanges between two legs as described by the soft anomalous dimension matrix (left), and possible exchanges as encoded in the evolution operator (right). Thick lines indicate the momenta of the final state colour charges originating from the hard scattering indicated by the grey blob.

integration per pair of partons, where we assume that μ^2 is always less than the scale of the invariant cluster masses M_{ij}^2.

At one loop, the soft anomalous dimension matrix reads [101]

$$\mathbf{\Gamma}(\{p\},\mu^2) = \sum_{i\neq j}(-\mathbf{T}_i\cdot\mathbf{T}_j)\Gamma_{\text{cusp}}\ \ln\left(\frac{-s_{ij}}{\mu^2}\right) + \sum_i \gamma_i, \tag{5.10}$$

where the sum runs over all pairs of parton legs, $s_{ij} = 2p_i\cdot p_j$ if both partons i and j are incoming or outgoing and $s_{ij} = -2p_i\cdot p_j$ if one is incoming and one is outgoing. The $\mathbf{T}_i\cdot\mathbf{T}_j$ are products of colour charge operators describing soft gluon exchange between any two legs i and j. For a review about colour charge operators, we refer to Ref. [36]. Γ_{cusp} is the so-called *cusp* anomalous dimension and in lowest order, can be written as $\Gamma_{\text{cusp}} = \alpha_s/4\pi$.

In general, we have to stick to a certain basis of colour structures to obtain a matrix representation of $\mathbf{\Gamma}$ such that an exponentiation can be performed, see Eq. (5.8). Since we are only interested in the evolution of the colour structures, we can neglect the non-cusp terms γ_i. They are diagonal in colour space and therefore have no effect on the resulting colour structure. If we assume that only final state massless partons are present, ($s_{ij} = 2p_i\cdot p_j$), the evolution operator becomes

$$\mathbf{U}\left(\{p\},\mu^2,\{M_{ij}^2\}\right) = \exp\left(-\sum_{i\neq j}\int_{\mu^2}^{M_{ij}^2}\frac{\mathrm{d}q^2}{q^2}(-\mathbf{T}_i\cdot\mathbf{T}_j)\Gamma_{\text{cusp}}\left(\ln\frac{2p_i\cdot p_j}{q^2} - i\pi\right)\right). \tag{5.11}$$

The evolution matrix $\mathbf{U}$ is a matrix exponential in colour space and describes soft-gluon exchanges between any two legs to all orders in the strong coupling. We have represented this in Fig. 5.1. The action of the evolution operator can be described as iterating colour reconnectors [98] which, once per action swap two indices of the permutation labelling the specific colour-flow and introduce longer transpositions when exponentiated. If the colour-flows represented through the basis tensor in Eq. (5.4) can be considered to represent physical colour singlet systems resulting in cluster assignments, then the evolution operator Eq. (5.11) can be seen as the basic object describing

the physics behind colour reconnection at the amplitude level. This assumption is the starting point for our model investigation.

In order to calculate the exponentiated matrix, which governs the evolution in colour space and to arrive at a formulation which can be implemented in form of a Monte Carlo model, we first have to determine the colour charge products $\mathbf{T}_i \cdot \mathbf{T}_j$. The action of a single colour charge operator, $\mathbf{T}_i$, where i denotes the label of the leg, corresponds to the emission of a gluon from leg i. This results in a higher dimension in colour space. The colour charge emission operator associated to each leg i can be conveniently written as [98]

$$\mathbf{T}_i = \lambda_i \mathbf{t}_{c_i} - \bar{\lambda}_i \mathbf{t}_{\bar{c}_i} - \frac{1}{N_c}(\lambda_i - \bar{\lambda}_i)\mathbf{s}, \tag{5.12}$$

where c_i, $\bar{c}_i$ are the colour labels assigned to the leg i and $\lambda_i = T_R$, $\bar{\lambda}_i = 0$ for a quark, $\lambda_i = 0$, $\bar{\lambda}_i = T_R$ for an antiquark and $\lambda_i = \bar{\lambda}_i = T_R$ for a gluon. The factor T_R is the representation index, also called *Dynkin* index and chosen as $1/2$ in the fundamental representation and N_c in the adjoint representation. The colour line operators $\mathbf{t}_\alpha, \bar{\mathbf{t}}_{\bar{\alpha}}$ are defined through their action on the colour base states as [99]

$$\mathbf{t}_\alpha|\sigma\rangle = \mathbf{t}_\alpha \left| \begin{matrix} 1 & \cdots & \alpha & \cdots & n \\ \sigma(1) & \cdots & \sigma(\alpha) & \cdots & \sigma(n) \end{matrix} \right\rangle = \left| \begin{matrix} 1 & \cdots & \alpha & \cdots & n & n+1 \\ \sigma(1) & \cdots & \sigma(n+1) & \cdots & \sigma(n) & \sigma(\alpha) \end{matrix} \right\rangle, \tag{5.13}$$

and

$$\bar{\mathbf{t}}_{\bar{\alpha}}|\sigma\rangle = \bar{\mathbf{t}}_{\bar{\alpha}} \left| \begin{matrix} 1 & \cdots & \alpha & \cdots & n \\ \sigma(1) & \cdots & \sigma(\alpha) & \cdots & \sigma(n) \end{matrix} \right\rangle = \left| \begin{matrix} 1 & \cdots & \sigma^{-1}(\alpha) & \cdots & n & n+1 \\ \sigma(1) & \cdots & \sigma(n+1) & \cdots & \sigma(n) & \sigma(\alpha) \end{matrix} \right\rangle. \tag{5.14}$$

The singlet operator $\mathbf{s}$ corresponds to the emission of a singlet gluon from a given leg. Since we work in the colour-flow basis the singlet operator will just add an additional colour line which is connected to itself

$$\mathbf{s}|\sigma\rangle = \mathbf{s} \left| \begin{matrix} 1 & \cdots & n \\ \sigma(1) & \cdots & \sigma(n) \end{matrix} \right\rangle = \left| \begin{matrix} 1 & \cdots & n & n+1 \\ \sigma(1) & \cdots & \sigma(n) & \sigma(n+1) \end{matrix} \right\rangle. \tag{5.15}$$

In order to calculate the products of colour line operators on a given colour flow, $\mathbf{t}_i \cdot \mathbf{t}_j$ which corresponds to the emission of a soft gluon from leg i and the absorption of the soft gluon from leg j, Eq. (5.12) proves to be helpful. The product of colour line operators, therefore the products of the colour charge operators $\mathbf{T}_i \cdot \mathbf{T}_j$ correspond to an exchange in colour space between the legs i and j. We calculate the matrix elements of the colour charge products following the description in Ref. [99] with the following formula

$$\begin{aligned}[\tau|\mathbf{T}_i \cdot \mathbf{T}_j|\sigma\rangle = -N_c\delta_{\tau,\sigma}[&\lambda_i\bar{\lambda}_j\delta_{\sigma(c_i),\bar{c}_j} + \bar{\lambda}_i\lambda_j\delta_{\sigma(c_j),\bar{c}_i} + \frac{1}{N_c^2}(\lambda_i - \bar{\lambda}_i)(\lambda_j - \bar{\lambda}_j)] \\ &+ \sum_{(ab)} \delta_{\tau_{(ab)},\sigma}\big(\lambda_i\lambda_j\delta_{(ab),(\sigma(c_i)\sigma(c_j))} + \bar{\lambda}_i\bar{\lambda}_j\delta_{(ab),(\bar{c}_i\bar{c}_j)} \\ &\qquad - \lambda_i\bar{\lambda}_j\delta_{(ab),(\sigma(c_i)\bar{c}_j)} - \bar{\lambda}_i\lambda_j\delta_{(ab),(\sigma(c_j)\bar{c}_i)}\big),\end{aligned} \tag{5.16}$$

for $i \neq j$. In this formula $\delta_{\tau\sigma}$ is zero if the permutations τ and σ are not equal, (ab) represents an ordered pair ($(ab) = ba$ if $a > b$) and $\tau_{(ab)}$ denotes swapping the elements a and b in the permutation τ. Since the colour-flow basis is non-orthogonal we have to introduce a dual basis vector $[\tau|$ [99], which is defined such that

$$\sum_\alpha |\alpha\rangle[\alpha| = \sum_\alpha |\alpha]\langle\alpha| = \mathbf{1}, \tag{5.17}$$

and

$$\langle\alpha|\beta] = [\alpha|\beta\rangle = \delta_{\alpha\beta}. \tag{5.18}$$

From Eq. (5.16) we can calculate all $\mathbf{T}_i \cdot \mathbf{T}_j$ matrix elements in the colour-flow basis.

5.3 Outline of the General Algorithm

The cluster configuration we obtain from a pre-confining parton evolution as discussed in Sec. 5.1 can be seen as driven by a cross-section which results from an amplitude $\mathcal{H}$ that has been dominated by a specific colour structure $|\sigma\rangle$. The colour structure results in the assignment of colour connected quark-antiquark pairs and therefore clusters in the pre-confined final state,

$$\mathrm{d}\sigma \sim |\mathcal{H}(\{p\}, Q^2, \{M_{ij}^2\})|^2, \qquad |\mathcal{H}(\{p\}, Q^2, \{M_{ij}^2\})\rangle \approx \mathcal{H}_\sigma(\{p\}, Q^2, \{M_{ij}^2\})|\sigma\rangle \ . \tag{5.19}$$

We further assume that the logarithms of Q^2/M_{ij}^2 have been summed by the parton shower evolution with $M_{ij}^2 \sim 2p_i \cdot p_j \sim Q_0^2$. This corresponds to a veto of radiation off dipoles with masses around the shower infrared cutoff Q_0^2. We view the initial step of colour reconnection as an evolution in colour space down to scales of order μ^2 below the initial cluster masses and the parton shower infrared cutoff. We then use

$$\mathbf{U}\left(\{p\}, \mu^2, \{M_{ij}^2\}\right) = \exp\left(\sum_{i\neq j} \mathbf{T}_i \cdot \mathbf{T}_j \frac{\alpha_s}{2\pi}\left(\frac{1}{2}\ln^2\frac{M_{ij}^2}{\mu^2} - i\pi \ln\frac{M_{ij}^2}{\mu^2}\right)\right), \tag{5.20}$$

as derived in Sec. 5.2 as an ansatz for the evolution of a given colour structure, where the term $-i\pi\ln(M_{ij}^2/\mu^2)$ is the so-called Coulomb term.

The starting point for the colour reconnection of a certain cluster configuration which is represented through a colour structure $|\sigma\rangle$ is to consider the overlap between the evolved amplitude, $\mathbf{U}\left(\{p\}, \mu^2, \{M_{ij}^2\}\right)|\sigma\rangle$, and a new colour structure $|\tau\rangle$. This allows us to define a reconnection amplitude

$$\mathcal{A}_{\sigma\to\tau} = \langle\tau|\mathbf{U}\left(\{p\}, \mu^2, \{M_{ij}^2\}\right)|\sigma\rangle. \tag{5.21}$$

From these amplitudes we can calculate probabilities to end up in all possible colour structures according to

$$P_{\sigma\to\tau} = \frac{|\mathcal{A}_{\sigma\to\tau}|^2}{\sum_\rho |\mathcal{A}_{\sigma\to\rho}|^2}\ , \tag{5.22}$$

where ρ runs over the possible colour-flows.

5.3.1 Example of two-cluster evolution

In this section we show the colour-flow evolution for the simple case of a two cluster system. For a two cluster system which consists of two quarks $(1, 2)$ and two antiquarks $(\bar{1}, \bar{2})$ there are two possible colour connections $(1\bar{1}, 2\bar{2})$ and $(1\bar{2}, 2\bar{1})$ each corresponding to a different cluster configuration resulting in different invariant cluster masses M_{ij}. In the colour-flow basis we represent the given colour-flow for a two cluster system in the basis state as

$$|\sigma\rangle = \left| \begin{matrix} 1 & 2 \\ \bar{1} & \bar{2} \end{matrix} \right\rangle = |12\rangle = \delta^{\alpha_1}_{\alpha_{\bar{1}}} \delta^{\alpha_2}_{\alpha_{\bar{2}}}, \tag{5.23}$$

where we have introduced the shorthand notation $|12\rangle$, which in our notation means that quark 1 is colour connected to antiquark $\bar{1}$ and quark 2 is colour connected to antiquark $\bar{2}$ resulting in the invariant cluster masses $M^2_{1\bar{1}}$ and $M^2_{2\bar{2}}$. It is furthermore helpful to represent the colour states as products of Kronecker deltas as shown in Eq. (5.23). For the exponential of the evolution operator, Eq. (5.20), that acts on a given colour-flow $|\sigma\rangle$ we can write

$$\Gamma = \Gamma_{1\bar{1}} \mathbf{T}_1 \mathbf{T}_{\bar{1}} + \Gamma_{12} \mathbf{T}_1 \mathbf{T}_2 + \Gamma_{1\bar{2}} \mathbf{T}_1 \mathbf{T}_{\bar{2}} + \Gamma_{\bar{1}2} \mathbf{T}_{\bar{1}} \mathbf{T}_2 + \Gamma_{\bar{1}\bar{2}} \mathbf{T}_{\bar{1}} \mathbf{T}_{\bar{2}} + \Gamma_{\bar{2}2} \mathbf{T}_{\bar{2}} \mathbf{T}_2, \tag{5.24}$$

where

$$\Gamma_{ij} = \frac{\alpha_s}{2\pi} \left(\frac{1}{2} \ln^2 \frac{M^2_{ij}}{\mu^2} - i\pi \ln \frac{M^2_{ij}}{\mu^2} \right). \tag{5.25}$$

The matrices of the products of colour charge operators can be calculated according to Eq. (5.16) and read

$$\mathbf{T}_i \cdot \mathbf{T}_{\bar{i}} = - \begin{pmatrix} N_c & 1 \\ 0 & 0 \end{pmatrix}, \; \mathbf{T}_i \cdot \mathbf{T}_{\bar{j}} = - \begin{pmatrix} 0 & 0 \\ 1 & N_c \end{pmatrix}, \; \mathbf{T}_i \cdot \mathbf{T}_j = - \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{5.26}$$

for the two cluster case. We stick to the matrix representation such that an exponentiation can be performed. From Eq. (5.24) in combination with the matrices Eq. (5.26) follows an explicit expression for the soft anomalous dimension matrix for the two cluster case which can then be exponentiated

$$\mathbf{\Gamma} = - \begin{pmatrix} -N_c(\Gamma_{1\bar{1}} + \Gamma_{2\bar{2}}) & \Gamma_{12} + \Gamma_{\bar{1}\bar{2}} - (\Gamma_{1\bar{1}} + \Gamma_{\bar{2}\bar{2}}) \\ \Gamma_{12} + \Gamma_{\bar{1}\bar{2}} - (\Gamma_{1\bar{2}} + \Gamma_{2\bar{1}}) & -N_c(\Gamma_{1\bar{2}} + \Gamma_{2\bar{1}}) \end{pmatrix}. \tag{5.27}$$

In this example we neglect the Coulomb term for simplicity but will include it in the later numerical simulations for the general case of up to 5 clusters. By neglecting the Coulomb term we obtain a compact form for the Γ_{ij} in terms of invariant cluster masses and the partons four-momenta

$$\Gamma_{ij} = \frac{\alpha_s}{2\pi} \Big[\frac{1}{2} \ln^2 \frac{M^2_{ij}}{\mu^2} \Big] = \frac{\alpha_s}{2\pi} \Big[\frac{1}{2} \ln^2 \frac{2 p_i \cdot p_j}{\mu^2} \Big]. \tag{5.28}$$

For the simple case of two-cluster evolution, the exponential of a general 2×2 matrix can be written down explicitly. Although already for this case, the exponential is quite complicated. For the later numerical studies the matrix exponential is performed

numerically by using the Padé approximation [103]. The matrix exponential for the two cluster case with $N_c = 3$ reads

$$\mathbf{U} = e^{\mathbf{\Gamma}} = e^{-\frac{3}{2}(a+b)} \sinh\left(\frac{\sqrt{\Delta}}{2}\right) \times \begin{pmatrix} \coth\left(\frac{\sqrt{\Delta}}{2}\right) + \frac{3(b-a)}{\sqrt{\Delta}} & \frac{2(c-a)}{\sqrt{\Delta}} \\ \frac{2(c-b)}{\sqrt{\Delta}} & \coth\left(\frac{\sqrt{\Delta}}{2}\right) + \frac{3(a-b)}{\sqrt{\Delta}} \end{pmatrix}, \tag{5.29}$$

where we introduced the variables $a = \Gamma_{2\bar{2}} + \Gamma_{1\bar{1}}$, $b = \Gamma_{1\bar{2}} + \Gamma_{2\bar{1}}$, $c = \Gamma_{12} + \Gamma_{\bar{2}\bar{1}}$ and $\Delta = 9a^2 - 4c(a+b) - 14ab + 9b^2 + 4c^2$. Eq. (5.29) is the matrix form of Eq. (5.20) governing the evolution of a given colour-flow. We can now proceed to determine the reconnection probability when the initial colour-flow is given by $|\sigma\rangle = |12\rangle$. We represent the two possible colour-flows by vectors in colour space

$$|12\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \; |21\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{5.30}$$

The dimension of the vectors equals the number of possible colour-flows The action of the evolution operator $\mathbf{U}$ on an initial colour-flow $|\sigma\rangle = |12\rangle$ results in a mixing of colour structures,

$$|\tau\rangle = \mathbf{U}|12\rangle = U_{12}|12\rangle + U_{21}|21\rangle, \tag{5.31}$$

where U_{12} and U_{21} are the matrix elements of the evolution operator. We can now proceed to determine a reconnection probability by projecting the evolved colour structure onto the alternative colour-flow $|21\rangle$ when the initial colour-flow before the evolution was $|12\rangle$

$$P_{\text{Reco}} = \frac{|\langle 21|\tau\rangle|^2}{|\langle 21|\tau\rangle|^2 + |\langle 12|\tau\rangle|^2}. \tag{5.32}$$

Since we work in a non-orthogonal basis we need to calculate the inner products of the colour-flow basis tensors $\langle ij|\tau\rangle$. It is helpful to consider the general expression of the basis states in terms of Kronecker deltas which leads to

$$\langle 12|12\rangle = \delta^{\bar{\alpha}_1}_{\alpha_1}\delta^{\alpha_1}_{\bar{\alpha}_1} \cdot \delta^{\bar{\alpha}_2}_{\alpha_2}\delta^{\alpha_2}_{\bar{\alpha}_2} = N_c^2, \tag{5.33}$$

where the indices run over the number of colours N_c (in our case $N_c = 3$). For $\langle 21|12\rangle$ follows

$$\langle 21|12\rangle = N_c. \tag{5.34}$$

This can be generalized to m clusters ($m!$ possible colour-flows) which leads to a general formula for the inner products

$$\langle\sigma|\tau\rangle = \delta^{\alpha_{\tau(1)}}_{\bar{\alpha}_{\tau(\bar{1})}} \ldots \delta^{\alpha_{\tau(m)}}_{\bar{\alpha}_{\tau(\bar{m})}} \delta^{\bar{\alpha}_{\sigma(\bar{1})}}_{\alpha_{\sigma(1)}} \ldots \delta^{\bar{\alpha}_{\sigma(\bar{m})}}_{\alpha_{\sigma(m)}} = N_c^{m-n}, \tag{5.35}$$

where n is the number of transpositions by which the two colour states $\langle\sigma|$ and $|\tau\rangle$ differ and N_c is the number of colours.

The reconnection probability P_{Reco} can then be expressed in terms of the matrix elements of the evolution matrix

$$P_{\text{Reco}} = \frac{|U_{11} + N_c U_{21}|^2}{|N_c U_{11} + U_{21}|^2 + |U_{11} + N_c U_{21}|^2}. \tag{5.36}$$

We will now proceed to explore the analytic properties of the probability to end up in the alternative colour-flow after the evolution. For minimal and maximal mixing in the evolution, the matrix elements of the evolution operator correspond to $U_{21} = 0$ and $U_{11} = 0$. The reconnection probabilities calculated from Eq. (5.36) with $N_c = 3$ are then $P_{\text{Reco}} = 1/10$, and $P_{\text{Reco}} = 9/10$ respectively. Most of the reconnection probability values are encountered for extremal cluster mass configurations: in the case when $U_{21} = 0$ the condition can be translated to the kinematical situation when the following equality is fulfilled

$$U_{21} = 0 \Leftrightarrow (\Gamma_{12} + \Gamma_{2\bar{1}} = \Gamma_{1\bar{2}} + \Gamma_{2\bar{1}})\,. \tag{5.37}$$

On the other hand, in the case $U_{11} = 0$ the value of $P_{\text{Reco}} = 9/10$ is obtained when

$$U_{11} = 0 = \sqrt{\Delta}\coth\left(\frac{\sqrt{\Delta}}{2}\right) + 3(b-a). \tag{5.38}$$

In the limit where $\sqrt{\Delta}$ is large[2], $\coth\left(\sqrt{\Delta}/2\right) \to 1$, and the condition above simplifies to

$$\frac{3(a-b)}{\sqrt{\Delta}} = 0 \Leftrightarrow (a-c)(c-b) = 0. \tag{5.39}$$

Since we assume that both U_{11} and U_{21} are not equal to 0 at the same time, the condition above is only fulfilled when

$$a - c = 0 \Leftrightarrow \Gamma_{2\bar{2}} + \Gamma_{1\bar{1}} = \Gamma_{12} + \Gamma_{\bar{2}\bar{1}}, \tag{5.40}$$

which is consistent with the numerical results presented in Sec. 5.5. In the next step we investigate the rapidity dependence of the reconnection probability. We work in the centre-of-mass frame of the particles 1 and $\bar{1}$ in which their momenta are in a back-to-back configuration. They are given by

$$p_1 = \frac{1}{2}(M_{1\bar{1}}, 0, 0, M_{1\bar{1}}), \tag{5.41}$$

$$p_{\bar{1}} = \frac{1}{2}(M_{1\bar{1}}, 0, 0, -M_{1\bar{1}}). \tag{5.42}$$

We express the four-momenta of the other two particles $i = 2, \bar{2}$ in dependence of their rapidities and azimuthal angles with respect to the quark axis of p_1

$$p_i = p_{Ti}(\cosh y_i, \sin\phi_i, \cos\phi_i, \sinh y_i). \tag{5.43}$$

[2]Already for $\Delta = 40$ the value of $\coth\left(\frac{\sqrt{\Delta}}{2}\right) = 1.00359$ and for $\Delta = 100$ it is equal to 1.00009.

The products then obtain the simple form:

$$p_1 \cdot p_{\bar{1}} = \frac{1}{2} M_{1\bar{1}}^2, \tag{5.44}$$

$$p_1 \cdot p_i = \frac{1}{2} M_{1i} p_{Ti} e^{-yi}, \tag{5.45}$$

$$p_{\bar{1}} \cdot p_i = \frac{1}{2} M_{\bar{1}i} p_{Ti} e^{yi}, \tag{5.46}$$

$$p_2 \cdot p_{\bar{2}} = 2 p_{T2} p_{T\bar{2}} (\cosh \Delta y_{2\bar{2}} - \cos \Delta \phi_{2\bar{2}}) \ , \tag{5.47}$$

and the condition for the minimal mixing from Eq. (5.37) can be rewritten as

$$\ln \left(\frac{1}{4} \frac{M_{1\bar{1}}^2}{\mu^2} \frac{p_{T2}^2}{\mu^2} \right) \ln \left(e^{-2y_2} \right) = \ln \left(\frac{1}{4} \frac{M_{1\bar{1}}^2}{\mu^2} \frac{p_{T\bar{2}}^2}{\mu^2} \right) \ln \left(e^{-2y_{\bar{2}}} \right). \tag{5.48}$$

When $p_{T2} \sim p_{T\bar{2}}$ and $\Delta Y = y_{\bar{2}} - y_2 \sim 0$, meaning that the initial partons in the clusters have similar transverse momenta and are close in rapidity, the reconnection probability Eq. (5.22) is minimal. Therefore, more likely will be reconnections when the ΔY values for the quark-antiquark pairs of the original clusters are large which is also confirmed numerically in Fig. 5.5. It is interesting to see the transverse momentum dependence of the result from Eq. (5.48). The impact of a transverse momentum dependence of the cluster constituents on the reconnection probability has been studied in Ref. [69].

5.4 Baryonic Reconnections

The concept of a baryonic type cluster has proven to be essential in improving the description of flavour observables at hadron colliders. Within the framework of perturbatively inspired colour reconnection we can accommodate for such reconnections provided there are at least 3 clusters present. It is then possible to associate a baryon-antibaryon pair to a colour-flow which has been anti-symmetrized in three fundamental and three anti-fundamental indices, in the following denoted by

$$|B_{ijk}\rangle = \frac{1}{N_B} \epsilon^{ijk} \epsilon_{\bar{i}\bar{j}\bar{k}} =$$
$$\frac{1}{N_B} \left(\left| \begin{matrix} i & j & k \\ \bar{i} & \bar{j} & \bar{k} \end{matrix} \right\rangle + \left| \begin{matrix} j & k & i \\ \bar{i} & \bar{j} & \bar{k} \end{matrix} \right\rangle + \left| \begin{matrix} k & i & j \\ \bar{i} & \bar{j} & \bar{k} \end{matrix} \right\rangle - \left| \begin{matrix} j & i & k \\ \bar{i} & \bar{j} & \bar{k} \end{matrix} \right\rangle - \left| \begin{matrix} i & k & j \\ \bar{i} & \bar{j} & \bar{k} \end{matrix} \right\rangle - \left| \begin{matrix} k & j & i \\ \bar{i} & \bar{j} & \bar{k} \end{matrix} \right\rangle \right). \tag{5.49}$$

The normalisation constant N_B is taken to reproduce the normalisation of a single mesonic configuration,

$$\langle B_{ijk} | B_{ijk} \rangle = N_c^3 \ , \qquad N_B^2 = 3! \left(1 - \frac{3}{N_c} + \frac{2}{N_c^2} \right) = \frac{4}{3}. \tag{5.50}$$

This allows us to define a *baryonic reconnection amplitude*

$$\mathcal{A}_{\tau \to B_{ijk} \otimes \tilde{\sigma}_{ijk}} = \langle B_{ijk} | \otimes \langle \tilde{\sigma}_{ijk} | \mathbf{U} \left(\{p\}, \mu^2, \{M_{ij}^2\} \right) | \tau \rangle, \tag{5.51}$$

where $\tilde{\sigma}_{ijk}$ denotes the permutation with the colour and anticolour indices corresponding to the baryonic system removed,

$$|\tilde{\sigma}_{ijk}\rangle = \left| \begin{array}{cccc} 1 & \cdots & n & \setminus\, i,j,k \\ \sigma(1) & \cdots & \sigma(n) & \setminus\, \bar{i},\bar{j},\bar{k} \end{array} \right\rangle . \tag{5.52}$$

The generalised reconnection probability is then[3]

$$P_{\tau\to\sigma} = \frac{|\mathcal{A}_{\tau\to\sigma}|^2}{\mathcal{N}_\tau} , \qquad P_{\tau\to B_{ijk}\otimes\tilde{\sigma}_{ijk}} = \frac{|\mathcal{A}_{\tau\to B_{ijk}\otimes\tilde{\sigma}_{ijk}}|^2}{\mathcal{N}_\tau}, \tag{5.53}$$

with

$$\mathcal{N}_\tau = \sum_\rho |\mathcal{A}_{\tau\to\rho}|^2 + \sum_\rho \sum_{i<j<k} |\mathcal{A}_{\tau\to B_{ijk}\otimes\tilde{\rho}_{ijk}}|^2. \tag{5.54}$$

We also consider the possibility of evolving an already existing baryon, for which we introduce *unbaryonizing* reconnection amplitudes

$$\mathcal{A}_{B_{ijk}\otimes\tilde{\sigma}_{ijk}\to\tau} = \langle\tau|\mathbf{U}\left(\{p\},\mu^2,\{M_{ij}^2\}\right)|B_{ijk}\rangle\otimes|\tilde{\sigma}_{ijk}\rangle. \tag{5.55}$$

These allow us to quantify how relevant such an evolution step would be for a high-mass baryonic system, which would not have entered the reconnection dynamics any more as in the case of the model considered in Ch. 4.

5.5 Numerical Results

A significant difference between this work and other colour reconnection models [7, 72] is that we do not directly compare clusters and then choose a configuration that would leave us with pre-specified properties such as a smaller invariant cluster mass. Since we calculate the probabilities to evolve into different colour structures we first show that our approach leads to reasonable results compatible with the effects of conventional colour reconnection algorithms. In order to analyse the effect of the colour reconnection we mostly compare kinematic variables associated to the clusters before and after reconnection. We first consider *mesonic* reconnections and later proceed to include *baryonic* reconnections. We generate the kinematics of the clusters using the RAMBO method [104], and a variation of the Jadach algorithm [105], which was used for the UA5 model [106]. While RAMBO is performing a flat phase space population, including a cluster configuration which would not be expected from a pre-confining shower evolution, the UA5 model provides us already with a physical mass spectrum.

5.5.1 Mesonic reconnections

A known issue which concerns the modelling of LHC events is that it is a priori not clear how to correlate the colour connections between different scattering centres of multiple parton interactions. The clusters emerging from these interactions are in general too heavy as they consist of quark-antiquark pairs which are not close in momentum space. In this case colour reconnection models are used to restore the notion of a colour pre-confined state leading to a shift towards smaller invariant cluster masses. In Fig. 5.2 we show the invariant mass distribution for the evolution of 5 cluster systems before

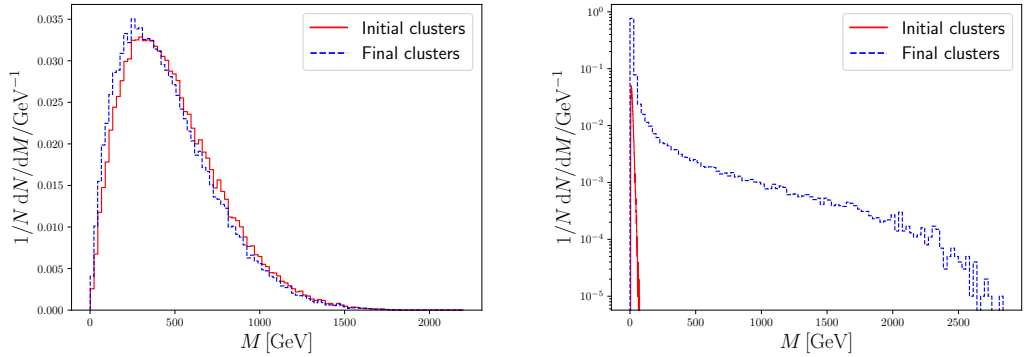

Figure 5.2: Invariant mass distribution of five-cluster evolution before and after mesonic colour reconnection for RAMBO(left) and UA5(right) kinematics. The quarks and antiquarks were generated at a centre-of-mass energy of $\sqrt{s} = 3000\,\mathrm{GeV}$. The parameter values are set to $\mu = 1\,\mathrm{GeV}$ and $\alpha_s = 0.118$.

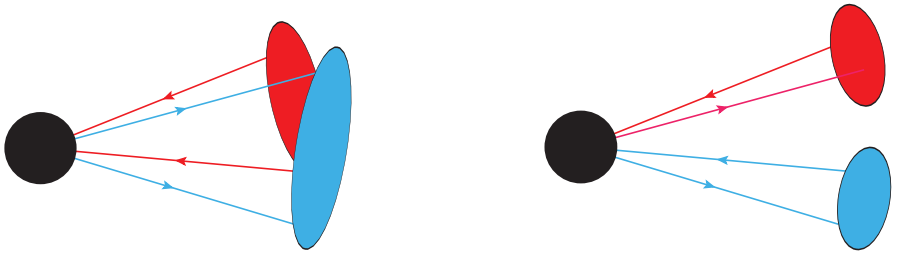

Figure 5.3: Sketch of the cluster configuration resulting from two alternative colour-flows from two-cluster evolution. The left figure shows two large overlapping clusters and the right figure shows a different colour-flow resulting in smaller clusters consisting of quark-antiquark pairs closer in momentum space. The black blob can be associated with an interaction or parts of an interaction that would lead to the shown configuration.

and after colour reconnection. While for the RAMBO kinematics the invariant mass distribution gets shifted towards smaller values, the clusters generated with the UA5 model already consist of quark-antiquark pairs close in momentum space, which leads to clusters with small invariant masses. Considering different colour flows will eventually connect quark-antiquark pairs well separated in rapidity leading to heavier clusters as seen in the right plot of Fig. 5.2.

When sampling the cluster kinematics with the UA5 model *but* enforcing random colour connections between the quarks and antiquarks for the initial configuration the resulting clusters are large and (in rapidity span) overlapping. A sketch of this configuration is shown in Fig. 5.3 for the case of two clusters. For this configuration our ansatz for colour reconnection again chooses colour-flows leading to a shift towards smaller values in terms of invariant cluster mass. This effect on the invariant mass spectrum can clearly be seen in Fig. 5.4, where we plot the logarithmic distribution of the invariant cluster masses. We conclude that our approach based on perturbative colour-flow evolution naturally prefers colour-flows leading to a configuration with smaller invariant

[3]Since it is the probability to end up with one specific colour-flow, the indices in the numerator are fixed.

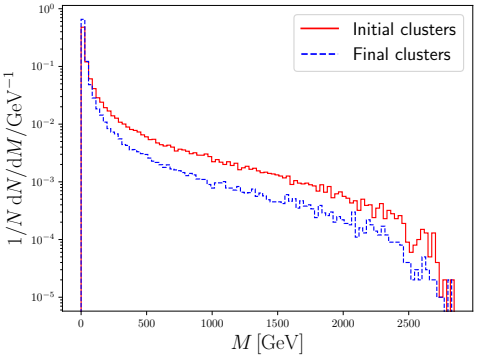

Figure 5.4: Invariant cluster mass distribution before and after colour reconnection for UA5 kinematics with random initial colour connections.

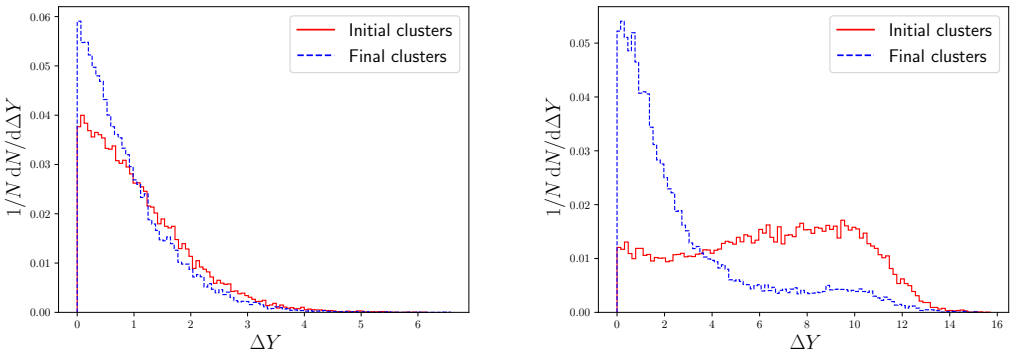

Figure 5.5: Histogram of the ΔY values for the quark-antiquark pairs of the original clusters that were reconnected and the ΔY value of the quark-antiquark pairs of the reconnected clusters. (a) RAMBO Phase space. (b) UA5 phase space with random colour connections.

cluster masses. We also stress here that the algorithm does not veto any colour-flows which would lead to an increase in terms of invariant cluster mass. To get an intuitive picture of what happens on the quark level, the rapidity difference ΔY between the quark-antiquark pairs which were participating in the reconnection process is shown in Fig. 5.5 for the RAMBO phase space and for the UA5 model with random initial colour connections. In both figures we see that colour-flows resulting in clusters consisting of quark-antiquark pairs which are closer in rapidity are clearly preferred. Again the effect is more pronounced for the UA5 model with random initial colour connections.

5.5.2 Baryonic reconnections

Within the context of our model a baryonic colour-flow can be introduced as explained in Sec. 5.4. In Fig. 5.6 the average baryonic reconnection probability for both phase space algorithms is shown. The first striking observation is that the probability rises with the number of clusters considered. The more clusters are present in an event, the more likely it is to find a candidate for baryonic reconnection. For RAMBO kinematics the initial colour configuration has no effect on the average reconnection probability. For the UA5 phase space it strongly depends on the initial colour configuration. Since

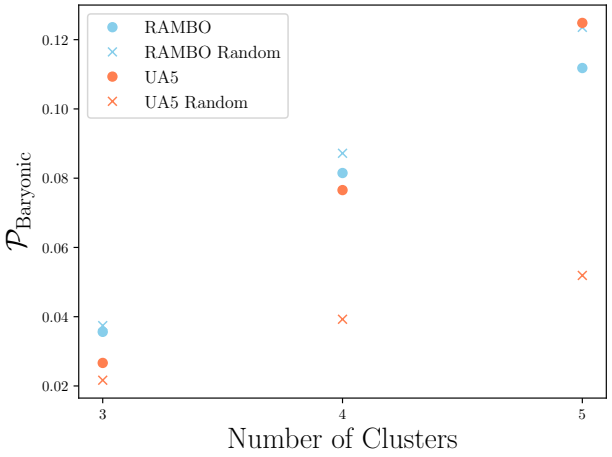

Figure 5.6: The average baryonic reconnection probabilities for the RAMBO and UA5 phase space with original and random initial colour connections. Depending on the phase space algorithm we employ to sample the initial configurations, the average baryonic reconnection probability ranges between $\approx 2\%$ and $\approx 12\%$.

the original UA5 cluster configuration already is in a state where the quarks are colour connected to their closest neighbours in phase space, the probability for reconnection into a different mesonic state is suppressed, increasing the probability to end up in a baryonic state. If the quarks are randomly connected, the average baryonic reconnection probability is suppressed since the probability for mesonic reconnection is high. Now we proceed to study the three, four and five-cluster evolution with the RAMBO phase space in detail. The distribution of the reconnection probabilities is shown in Fig. 5.7. The baryonic reconnection probability tends to prefer smaller values with a pronounced peak at zero. The tail towards higher values in the distribution might indicate some preferred kinematic configurations for the evolution into a baryonic state. With only one possible baryonic configuration, the three-cluster evolution is convenient to analyse and to extract a kinematic dependence.

In Fig. 5.8 the probability to evolve into a baryonic state with respect to the sum of average ΔR values of the quarks and antiquarks that would constitute a baryonic cluster, $(\langle \Delta R_B \rangle + \langle \Delta R_{\bar{B}} \rangle)/2$ is shown, where ΔR is defined as the distance between the constituent quarks in the $y - \phi$ plane

$$\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta y)^2}, \tag{5.56}$$

and we define $\langle \Delta R_{B,\bar{B}} \rangle$ as

$$\langle \Delta R_{B,\bar{B}} \rangle = (\Delta R_{12,\bar{1}\bar{2}} + \Delta R_{13,\bar{1}\bar{3}} + \Delta R_{23,\bar{2}\bar{3}})/3, \tag{5.57}$$

where the subscripts $(1, 2, 3)$, $(\bar{1}, \bar{2}, \bar{3})$, denote the quarks(antiquarks) inside the clusters of baryonic(anti-baryonic) type. The median shows a rising probability with smaller $(\langle \Delta R_B \rangle + \langle \Delta R_{\bar{B}} \rangle)/2$ values which indicates that the formation into a baryonic cluster is preferred if the three quarks *and* the three antiquarks are close together in ΔR space. We note that the baryonic and the anti-baryonic cluster can still be overlapping since we do not take the distance between them into account.

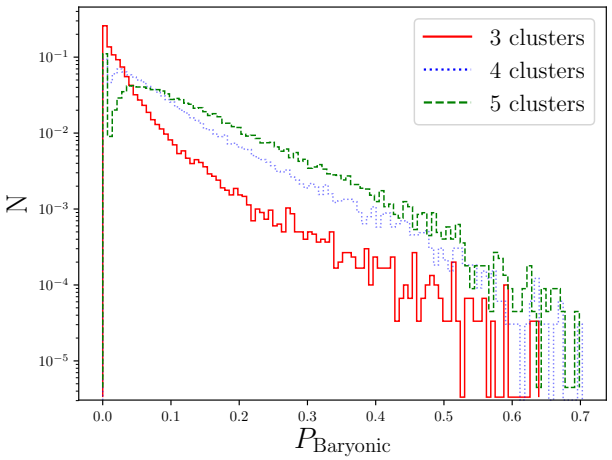

Figure 5.7: Histogram of baryonic reconnection probabilities for three, four and five-cluster evolution where the quark kinematics was sampled with the RAMBO phase space algorithm.

5.5.3 Unbaryonization

The colour reconnection algorithm presented in Ch. 4, that allows reconnection into a baryonic state is structured in a way, that once a baryonic cluster is formed, it is not considered for any further modifications. This clearly biases the reconnection procedure but has been necessary in order to cope with the rising complexity of many-cluster systems. In principle, a system could evolve into a baryonic state and then evolve again into a mesonic state which in turn lowers the amount of baryonic clusters occurring in an event. In the context of our model we can study the evolution back into a mesonic state by considering *unbaryonization* where we start the evolution with a baryonic configuration as the initial state and calculate the probability to evolve into a mesonic cluster configuration. In Fig. 5.9 we show the average probabilities for unbaryonization in terms of $(\langle \Delta Y_B \rangle + \langle \Delta Y_{\bar{B}} \rangle)/2$ and $(\langle \Delta R_B \rangle + \rangle \Delta R_{\bar{B}} \rangle)/2$, where $\langle \Delta Y_B \rangle = (|y_{q1} - y_{q2}| + |y_{q1} - y_{q3}| + |y_{q2} - y_{q3}|)/3$. This further adds to the intuitive picture that the probability for forming a baryonic cluster is high if the three quarks and the three antiquarks are close in momentum space. This also suggests that the colour field between quarks is enhanced if they do move in the same direction. In principle this could be used to decide whether a baryonic cluster should be kept or not. This allows for more flexibility and may introduce less bias in the reconnection algorithm.

5.5.4 Parameter variations and general findings

The ansatz for the colour-flow evolution Eq. (5.20) depends on the two parameters μ and α_s, as well as on the Coulomb term. The reconnection probability is a dynamic quantity. It strongly depends on the kinematics of the cluster constituents before and after reconnection and the parameter μ which can be viewed as a cutoff parameter of the colour-flow evolution in Eq. (5.11). In Fig. 5.10 we show the distribution of invariant

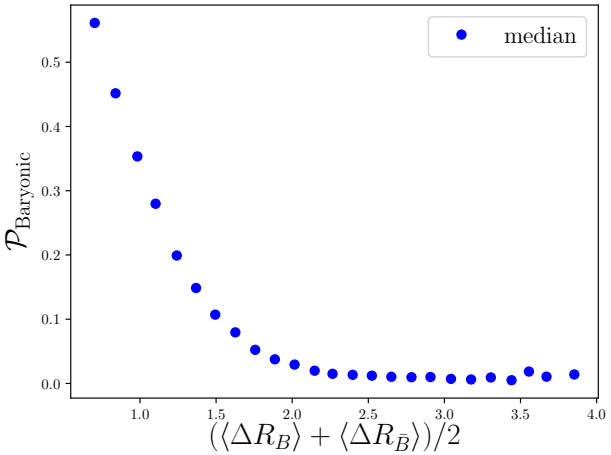

Figure 5.8: Median values for the baryonic reconnection probability with respect to $(\langle \Delta R_B \rangle + \langle \Delta R_{\bar{B}} \rangle)/2$ of the baryonic clusters for three-cluster evolution.

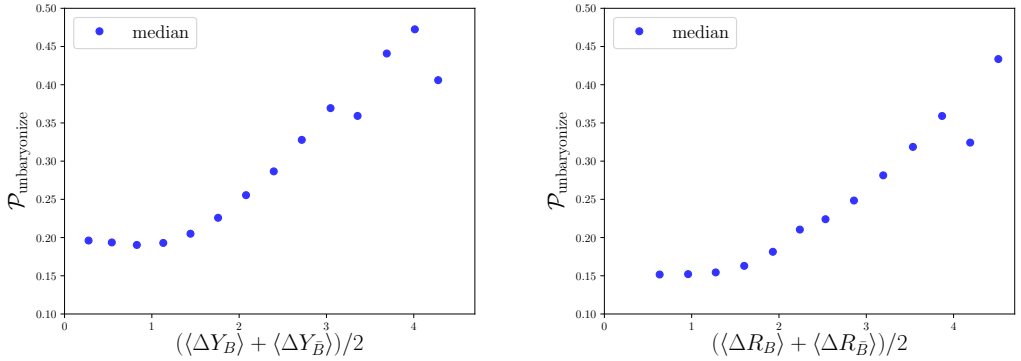

Figure 5.9: Average unbaryonization probabilities with respect to the average sum of $\langle \Delta Y \rangle$ and $\langle \Delta R \rangle$ of the constituent quarks of the baryonic clusters.

cluster masses for four-cluster evolution with two different values of $\mu = \{1, 0.01\}$ GeV and the corresponding colour length drop [72], which is defined as

$$\Delta_{\text{if}} = 1 - \frac{\lambda_{\text{final}}}{\lambda_{\text{initial}}}, \tag{5.58}$$

where λ_{initial} and λ_{final} denote the colour length before and after colour reconnection in an event which is defined as the sum of squared invariant cluster masses

$$\lambda = \sum_{i=1}^{N_{\text{cl}}} M_i^2. \tag{5.59}$$

In case of no colour reconnection $\lambda_{\text{initial}} \approx \lambda_{\text{final}}$ the colour length drop Δ_{if} approximately vanishes. If $\Delta_{\text{if}} \approx 1$, there is a significant change in λ which indicates a big effect due to colour reconnection. The kinematics of the four clusters were sampled with the RAMBO method. In order to have more physical cluster masses we sample them with a centre-of-mass energy of $\sqrt{s} = 10\,\text{GeV}$ which is closer to the cluster mass spectrum at the end of a typical shower evolution. Comparing the four-cluster evolution with the

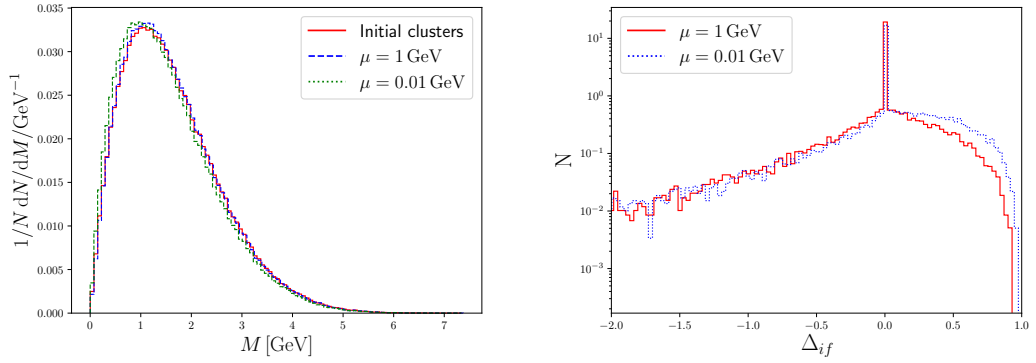

Figure 5.10: Distribution of invariant cluster masses before and after colour reconnection and colour length drop for different cut-off values of $\mu = \{1, 0.01\}$ GeV.

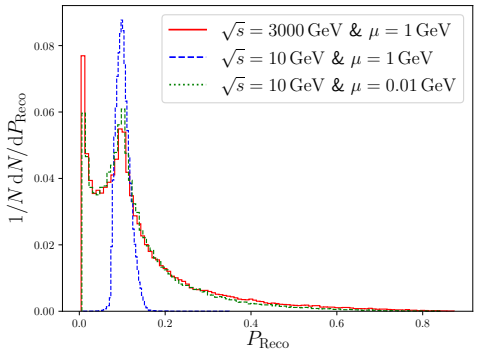

Figure 5.11: Reconnection probability for the two-cluster evolution for $\sqrt{s} = 3000\,\mathrm{GeV}$ and $\sqrt{s} = 10\,\mathrm{GeV}$.

different values for μ we see that $\mu = 0.01$ GeV leads to a small shift towards smaller invariant cluster masses. The distribution of Δ_{if} peaks at zero and is then distributed towards the positive and negative region. Negative values of Δ_{if} are also possible since we do not veto any colour-flows which would result in larger invariant cluster masses. The impact of the colour reconnection algorithm is more pronounced for $\mu = 0.01$, shifting the distribution of invariant cluster masses towards smaller values which can also be seen for Δ_{if}. If $\mu \to 0$ the colour-flow evolves into a configuration which results in preferably low invariant cluster masses.

In Sec. 5.5.1 we showed the effect of the algorithm on a relatively unphysical distribution of cluster masses as a simple proof of concept, that our ansatz and our algorithm indeed produce reasonable results. In this section we study the evolution of colour-flow and the behaviour of the model at different centre-of-mass energies $\sqrt{s}$ with the RAMBO method where we compare $\sqrt{s} = 3000\,\mathrm{GeV}$ with a more physical centre-of-mass energy of $\sqrt{s} = 10\,\mathrm{GeV}$, probing a spectrum of smaller clusters. We show the reconnection probability P_{Reco} for the case of two-cluster evolution for the two different centre-of-mass energies with different μ values in Fig. 5.11. While at $\sqrt{s} = 3000\,\mathrm{GeV}$ and $\mu = 1$ GeV the reconnection probabilities cover the whole range from zero to one, the reconnection probabilities are narrowly distributed around $P_{\mathrm{Reco}} \approx 0.1$ for $\sqrt{s} = 10\,\mathrm{GeV}$ and $\mu = 1\,\mathrm{GeV}$. This can be countered by reducing the μ parameter such that the ratio

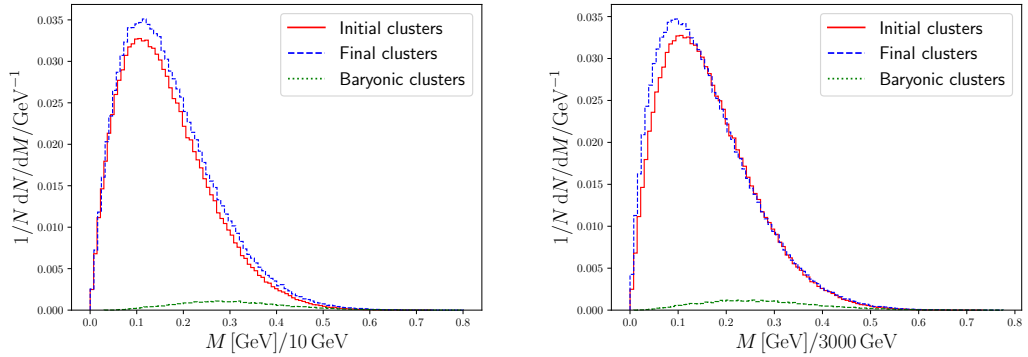

Figure 5.12: Distribution of invariant cluster masses for $\sqrt{s} = 10$ GeV and $\sqrt{s} = 3000$ GeV divided by their respective $\sqrt{s}$.

in the logarithm has the same value as for higher centre-of-mass energies as is also shown in Fig. 5.11 for the two-cluster evolution with $\sqrt{s} = 10\,\mathrm{GeV}$ and $\mu = 0.01\,\mathrm{GeV}$.

With a smaller μ parameter the reconnection probabilities start to cover the whole range. We observe the same behaviour for baryonic reconnections. If μ is small, the amplitude in colour space will continue to evolve down to a much lower scale. In the end, this results in a colour-flow with preferably small invariant cluster masses. In a full model the parameter μ could be tuned to data in order to define the cutoff at which the evolution is bound to stop. This could be used to verify if the amplitude does indeed favour a colour-flow resulting in small invariant cluster masses or one which allows for fluctuations towards larger invariant cluster masses. In Fig. 5.12 we show the mass distributions of 4 clusters for $\sqrt{s} = 10\,\mathrm{GeV}$ and $\sqrt{s} = 3000\,\mathrm{GeV}$ divided by the respective centre-of-mass energy with the possibility to produce baryonic clusters. The distributions of the reconnected clusters are similar, although for $\sqrt{s} = 10\,\mathrm{GeV}$ there is less room to evolve into a state of smaller cluster masses since we used $\mu = 1\,\mathrm{GeV}$. The mass distribution of baryonic clusters is also shifted between the two centre-of-mass energies. Since the soft anomalous dimension matrix only depends on the ratio of the invariant cluster masses and the cut off parameter μ, it is possible to find an energy independent prescription which should lead to the same distribution of cluster masses. A continuation of the arguments of the logarithms $\ln(M^2_{\alpha\beta}/\mu^2) \to \ln(M^2_{\alpha\beta}/\mu^2 + 1)$ has been considered but did not show any change in our findings and can as such be used to prevent numerical instabilities should very small invariant cluster masses be encountered in a full model. The strong coupling parameter α_s functions as a direct measure of the overall reconnection strength. We have chosen to reflect the strong coupling $\alpha_s = 0.118$ at a lower scale, but it should in practice also be considered a tunable parameter of the model.

Until now we neglected quark masses completely, which lead to a simplified form of the invariant cluster mass. Assuming the same quark (constituent) masses for light quarks (0.3 GeV for up and down quarks) , which are used in the cluster hadronization model of Herwig 7, only small effects were found in terms of the mass distribution and no sizeable effects on the reconnection probabilities. For heavy quarks ($m_b \sim 4 - 5\,\mathrm{GeV}$) more severe effects are expected but we leave this topic for a detailed study in the scope of a full model implementation. Neglecting the Coulomb term does not change our findings for the high-mass systems, while we see some effects for small-mass systems.

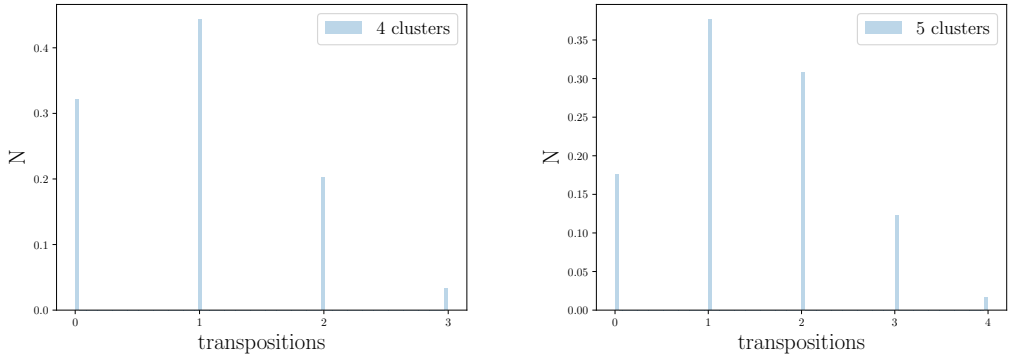

Figure 5.13: Number of transpositions between the initial state and the reconnected final state for four and five-cluster evolution.

5.6 Towards a Full Model

Due to its complexity, the approach followed in this chapter is limited to a small number of clusters and should be seen as a theoretical consideration to constrain the structure of colour reconnection models. It makes a full colour-flow evolution unfeasible to implement on the typically large systems of clusters encountered in a high-energy hadronic collisions. We will use the insights we gained from looking at the evolution of small systems to extrapolate a simplified model which could be suitable for implementation. A crucial aspect in this regard is to identify independently evolving subsystems. This would allow us to repeatedly evolve smaller subsystems of clusters within a larger ensemble.

The cluster configurations, being colour structures in the colour-flow basis, can be labelled by permutations. An important quantity is the minimum number of transpositions we require to transform the initial configuration into a reconnected final configuration. This number is directly related to the power of the number of colours N_c when evaluating the overlap between the two colour structures. The more transpositions are needed, the higher is the $1/N_c$ suppression. In Fig. 5.13 we show the number of transpositions for the evolution with four and five clusters where the phase space has been populated with the RAMBO method. For both cases we note the peak at one transposition which corresponds to a reconnection within a two cluster system, between the initial and final state. This indicates the existence of independently evolving subsystems where the contributions from the remaining clusters are suppressed. This suggests that colour reconnection is not simply a $1/N_c$ suppressed effect which would have indicated a much higher rate of non-reconnected systems, as well as a much steeper drop of the other reconnection dynamics with the number of transpositions. A simplified sketch, where the mixing with a well separated cluster system can be neglected is show in Fig. 5.14. Colour reconnection models implemented in event generators often rely on very simplified models in order to handle the complex structure of hadronic collisions. In the plain colour reconnection model Herwig 7, reconnections would only be accepted if they allow for a smaller mass configuration. They are accepted with a fixed probability which was inferred by tuning to MB data. While this approach has benefits in terms of efficiency and simplicity, and has shown to provide a reasonable description of data [72], it does not take into account the full kinematic dynamics and complexity of a hadronic event.

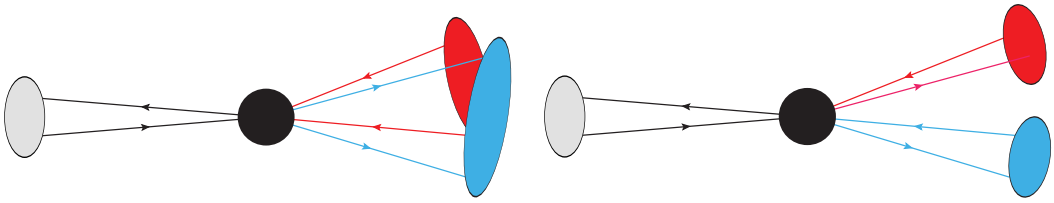

Figure 5.14: Sketch of a configuration where the two coloured clusters are assumed to be independently evolving subsystem. Gluon exchanges which would lead to colour connections between the clusters are expected to be highly suppressed. Left: Initial colour flow. Right: Colour flow after reconnection.

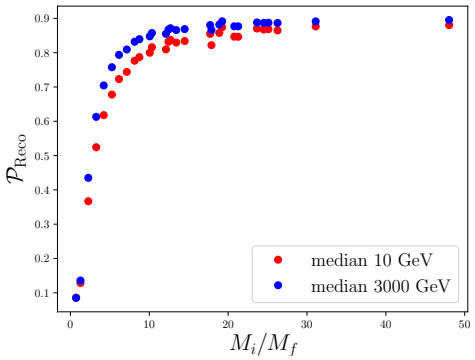

Figure 5.15: Parametrization of the reconnection probability for two-cluster evolution in terms of the ratio of invariant cluster masses M_i/M_f, where M_i is the sum of invariant cluster masses of the initial colour-flow and M_f is the sum of invariant cluster masses of the alternative colour-flow of a two cluster system.

In order to make contact with the plain colour reconnection model, which is based on a direct reduction of invariant cluster masses we consider the reconnection probability projected to a variable of the plain colour reconnection model, the ratio of the sum of cluster masses before and after reconnection. To do so we generate the kinematics of a two cluster system using the RAMBO algorithm with a centre-of-mass energy of $\sqrt{s} = 10\,\text{GeV}$ and with the parameter values $\mu = 0.01\,\text{GeV}$ and $\alpha_s = 0.118$. For comparison we also generate the kinematics with a higher centre-of-mass energy of $\sqrt{s} = 3000\,\text{GeV}$ and the parameter values $\mu = 1\,\text{GeV}$, $\alpha_s = 0.118$. In Fig. 5.15 we plot the median of the calculated reconnection probabilities over the ratio of the sum of invariant cluster masses which would result from the two possible colour-flows. The plain colour reconnection model would only have put a step function with no further kinematic dependency present in place. The reconnection probability saturates between $P_{\text{Reco}} \approx 0.8$ and $P_{\text{Reco}} \approx 0.9$ if reconnection would results in a smaller sum of invariant cluster masses. This limit is already obtained from our analytic studies in Sec. 5.3. We conclude that the dependence of the reconnection probability on the ratio of cluster masses is more dynamic than a simple step function as implemented in the plain colour reconnection model of Herwig 7. We suggest to use our findings as an input to improve the performance of colour reconnection models:

- With the finding of mostly independently evolving two cluster systems the analytically known reconnection probability from the evolution of a cluster systems can be directly used to improve the model assumption of these type of reconnections.
- The fact that also three cluster systems seem to be rather detached in a large ensemble can be used to supplement the baryon production mechanism as presented in Ch. 4 and Ref. [7] with a more dynamic reconnection probability.
- The baryonic reconnection mechanism and the possibility of un-connecting baryonic clusters in an evolution picture or statistical model [72].

All of these mechanisms are contained within our approach. With the analogue of the strong coupling α_s and the evolution cutoff μ it effectively contains two parameters. If our approach is also continued to small scales in a non-perturbative context, the number of colours N_c could be considered a tuning parameter as well. We finally note that, although we have essentially been considering the cluster hadronization model, similar dynamics could be implemented in a string picture.

5.7 Conclusion

We have studied to what extent the structure of perturbative colour evolution can be used as an input to improve or constrain existing colour reconnection models. We did consider the cluster hadronization model and in particular, we have analytically solved the evolution of a two-cluster system, as well as numerically studied the evolution of larger systems of up to five clusters. We have found that there is indeed a highly dynamic and non-trivial re-arrangement of colour structures already from a simple ansatz using a one-loop soft anomalous dimension, which confirms earlier work on geometrically inspired reconnection models.

The full evolution in colour space is, however, not feasible in a realistic model which needs to cope with tens to hundreds of clusters. However we have found indications that in the evolution of larger systems the bulk of the reconnection effects is isolated in small subsystems of two to three clusters which allows to build an iterative model. Our framework also allows to include a probability of reconnecting baryonic clusters into mesonic systems, an aspect which has not been considered in models of baryonic reconnection so far.

We stress the fact that our approach should be considered as a motivation for improved models of non-perturbative colour reconnection. It highlights that the perturbative mixing of colour flows, mediated through virtual soft gluon exchanges, should be considered an important ingredient in new approaches aiming to improve parton shower algorithms beyond the leading-colour level [99]. However until such algorithms are fully available, and the dynamics of hadronization is understood in this context, we postpone further aspects to future work and use the findings obtained here as an improved input to existing colour reconnection models.

CHAPTER 6

Kinematic Strangeness Production

The non-perturbative aspects of simulating LHC events remain an active area of research in light of the recent observations of ALICE [76] and CMS [75]. Signs of strangeness enhancement and collective effects in high-multiplicity events have inspired several phenomenological models, auch as interacting strings [86, 107], relativistic hydrodynamics [108], tweaks to the existing multiple parton interaction mechanisms [109] and colour reconnection models [7, 77].

The most successful hadronization models which try to reproduce the observed strangeness enhancement in high-multiplicity pp, pPb, $PbPb$ events are rooted in the physics of collectivity. The dense environment of these events leads to more complicated systems interacting with each other introducing non-trivial correlations in the final state. Heavy ion event generators typically prefer a hydrodynamic viewpoint, where the quark-gluon plasma acts as a perfect fluid, changing the dynamics of hadronization. High-energy pp event generators, on the other hand, tend to use sophisticated iterations of the more conventional proton-proton collision techniques.

In Ch. 4 we introduced a new model for colour reconnection, where baryonic clusters are allowed to be produced in a geometric fashion, to explain the results of Ref. [76]. The model was able to create heavier hadrons, and in particular more baryons. But in order to improve the description of the data, the non-perturbative gluon splitting mechanism was allowed to produce $s\bar{s}$ pairs as well as the default lighter species. However, the production weight was simply set to a flat number, tuned to MB events at the LHC. In this chapter, we will mainly focus on the fundamental mechanisms of strangeness production in cluster hadronization, namely the production rate of $s\bar{s}$ pairs during non-perturbative gluon splitting, cluster fission, and cluster decay. In doing so, we are taking the first steps towards a rework of strangeness production in the hadronization model of Herwig 7. A full model would also need to consider colour reconnection, since this rearranges the colour topology and thus the mass distribution inside an event, affecting the scaling that we are interested in studying.

We start the discussion with a review of the strangeness production mechanism in Herwig 7 in Sec. 6.1. We then perform two separate tunes to a number of light strange meson observables for LEP and LHC MB events in Sec. 6.2. We show that the current tuned strangeness production parameters are drastically different between the two collider types, and propose a mass-based scaling for the relevant production weights in Sec. 6.3, where we compare two different mass-like measures. In Sec. 6.4, we tune the new model for strangeness production and compare the results to the default model used in Herwig 7[4], as well as the default model in Pythia with the Monash tune [110]. We briefly summarize our work in Sec. 6.5.

6.1 The Strangeness Parameters of Herwig 7

The stages of hadronization which contribute to the overall strangeness content of the event are:

- non-perturbative gluon splitting,
- cluster fissioning,
- cluster decay.

The Herwig 7 parameters that control non-perturbative strangeness production are the gluon splitting weight - `SplitPwtSquark` and the cluster fission & decay weight - `PwtSquark`. In the default model, cluster fissioning and cluster decaying are controlled by the same parameter. The first step in the understanding of the different contributions is to disentangle cluster fission from cluster decay and introduce one additional parameter which controls the production of a $s\bar{s}$ pair during cluster fission - `FissionPwtSquark`. The decay parameter remains the same.

6.2 Tuning of the Existing Model

In this section we tune the parameters for strangeness production of the default Herwig 7 model first to LEP and then to LHC data. In order to understand the overall effects of strangeness production during different stages of the event generation, we keep all other hadronization parameters that were previously tuned to LEP data at their default values [3, 4]. In the first tune (TUNE1), we only consider the effects of the parameters that are *directly* responsible for strangeness production as explained in Sec. 6.1. In a second tuning attempt (TUNE2), we introduce a new parameter for the cluster fissioning stage (`FissionPwtSquark`). Tuning these 3 different parameters will allow us to study the phases of strangeness production during the event generation and enables us to study the differences between LEP and LHC.

6.2.1 LEP Tuning

For the tuning to LEP data, the following observables from ALEPH [111], DELPHI [112], SLD [113] and PDG hadron multiplicities [114], which represent a good description of event shapes and π, K multiplicities, were used with equal weights:

[4]The default model of Herwig 7 refers to Herwig 7.15 without baryonic colour reconnection and $g \to s\bar{s}$ gluon splitting

- Mean charged multiplicities for rapidities $|y| < 1.0$, $|y| < 1.5$ and $|y| < 2.0$,
- K^0 spectrum,
- Mean π^0 multiplicity,
- Mean $K_S + K_L$ multiplicity,
- Mean K^0 multiplicity,
- Mean π^+/π^- multiplicity,
- Mean K^+K^- multiplicity,
- Ratio (w.r.t $\pi^\pm$) of mean $K^\pm$ multiplicity,
- Ratio (w.r.t $\pi^\pm$) of mean K^0 multiplicity,
- $K^\pm$ scaled momentum.

The resulting parameter values for the two different tunes are listed in Tab.6.1.

LEP	Default	TUNE1	TUNE2
Gluon Splitting	–	0.24	0.19
Cluster Fission	0.66	0.53	0.690
Cluster Decay	0.66	0.53	0.685

Table 6.1: Results of the parameter values for strangeness production at the different stages of the event generation (LEP). In both default Herwig 7 and TUNE1, cluster fission and decay have the same parameter. In TUNE2, they are allowed to be different, but the tuning procedure returned almost equal values. In default Herwig 7, there is no $g \to s\bar{s}$ option.

While being able to describe all the considered LEP data on equally good footing as the default version of Herwig 7, we improve the simulation of the observables which were considered in the tuning procedure. TUNE2 gives better agreement with data, especially with respect to the $K^\pm$ multiplicity, highlighting the necessity to disentangle the cluster fission and cluster decay parameters. The corresponding plots are shown in Fig. 6.1, where we compare the default version with the two new tunes.

6.2.2 LHC Tuning

For the tuning to LHC data, we solely focus on identified particle distributions which were measured at ALICE [90] and CMS [75]. We limit the tuning to a center of mass energy of $\sqrt{s} = 7$ TeV due to the lack of suitable available Rivet analyses at higher energies. The following observables were considered in the tuning procedure with equal weights:

- $K^+ + K^-$ yield in INEL pp collisions at $\sqrt{s} = 7$ TeV in $|y| < 0.5$,

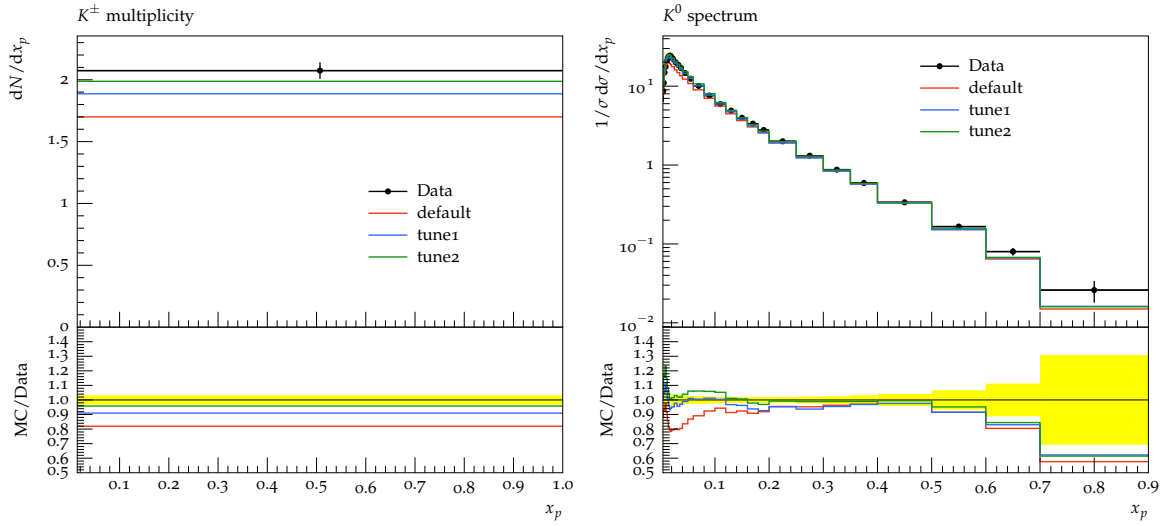

Figure 6.1: Measurement of $K^{\pm}$ multiplicities at SLD [113] and K^0 spectrum as measured at ALEPH [111] for $\sqrt{s} = 91.2$ GeV. We show a comparison between the default Herwig 7 model and the two different tunes.

- K/π in INEL pp collisions at $\sqrt{s} = 7$ TeV in $|y| < 0.5$,
- K^0_S rapidity distribution at $\sqrt{s} = 7$ TeV,
- K^0_S transverse momentum distribution at $\sqrt{s} = 7$ TeV.

The resulting parameter values are shown in Tab. 6.2. The outcome of the tuning procedure is shown in Fig. 6.2 for the p_T distribution of $K^+ + K^-$ yields and the K/π ratio.

LHC	Default	TUNE1	TUNE2
Gluon Splitting	–	0.95	0.95
Cluster Fission	0.66	0.05	0.02
Cluster Decay	0.66	0.05	0.25

Table 6.2: Results of the parameter values for strangeness production at the different stages of the event generation (LHC). In both default Herwig 7 and TUNE1, cluster fission and decay have the same parameter, while in TUNE2 they are allowed to be different. In default Herwig 7, there is no $g \to s\bar{s}$ option.

Similar to LEP, we observe that the re-tuning of the default model with the incorporation of an additional independent parameter at the cluster fission stage improves the description of the considered observables significantly.

6.2.3 Summary

The general approach in tuning a hadronization model is to tune the parameters to LEP data and then assume the model is able to describe LHC observables as well, since hadronization is expected to factorize and should not depend on the process involved.

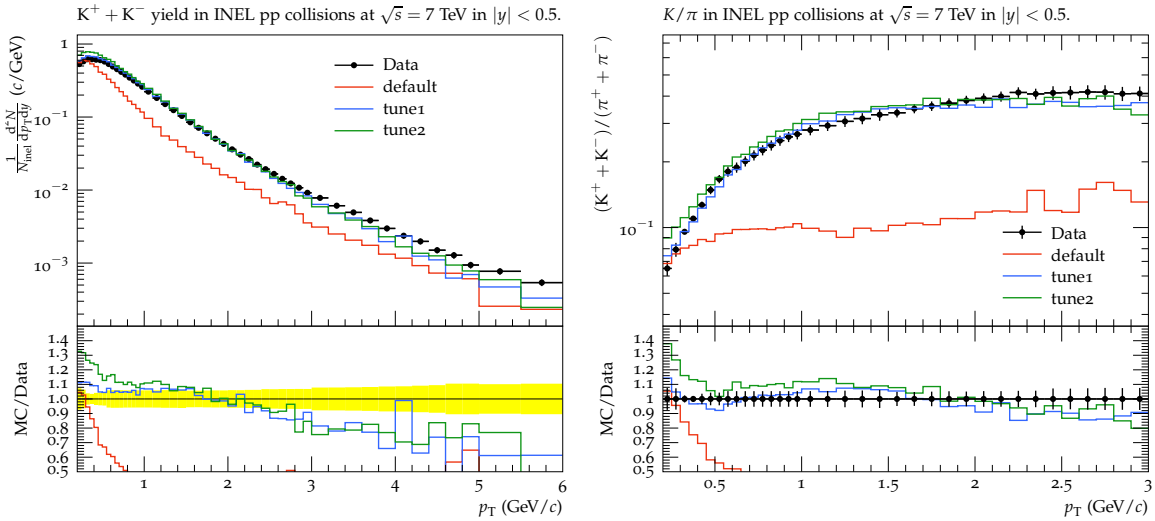

Figure 6.2: Transverse momenta spectra for $K^+ + K^-$ and K/π ratio as measured by ALICE [90] at $\sqrt{s} = 7$ TeV in the central rapidity region. We show a comparison between the default Herwig 7 model and the different tunes.

The main difference between LEP and LHC is the denser hadronic environment due to multiple parton interactions and therefore also the enhanced effect of colour reconnections on the distribution of final state particles. We believe that the probability to produce strangeness e.g. at the stage of non-perturbative gluon splitting should be a universal parameter and be independent of the considered process.

The data shows that different parameter values are preferred at LHC and LEP. We conclude that the approach to have a single valued probability is not suited for the description of both LHC and LEP observables. It may capture the average effect but it does not allow for fluctuations on an event-by-event basis. We tackle this issue by assuming that the rate at which strangeness is produced depends on the hadronic density of the immediate environment. This will be discussed in the next section.

6.3 Model for Kinematic Strangeness Production

As mentioned above, the various splitting probabilities and weights are flat numbers tuned to data, without any considerations for the topology of a given event. In order to have a more dynamic picture, where the splitting probabilities depend on the environment, we choose to scale the weights with respect to colour-singlet masses. The mass of a colour-singlet system at a given stage of hadronization then increases or decreases the probability for strangeness production.

As a simple starting point for mass-based scaling, we replace the flat weights in each of the steps mentioned in Sec. 6.1 with the functional form

$$w_s(m^2) = \exp\left(-\frac{m_0^2}{f(m^2)}\right), \qquad f(m^2) > 0, \tag{6.1}$$

which depends on the mass m of the colour-singlet system. In Eq. (6.1) m_0^2 is the characteristic mass scale for each phase, which is a tunable parameter and has distinctive values for each stage of the hadronization process. The function $f(m^2)$ is a mass-based

measure of the relevant colour-singlet system. We will use two different mass-measures: the invariant mass of a colour-singlet system, and the threshold production measure, λ. We discuss the difference in the two approaches in Sec. 6.3.3. In the following sections, we will use the invariant mass as an example. The weight, Eq. (6.1), is the weight for strangeness production, and is seen to be relative to the production weights of up and down quarks. In the limit of a heavy colour-singlet, the rate of producing strangeness will be the same as that of the lighter quarks, while in the low-mass limit, only the lighter quarks will be allowed to be produced. The appeal of an exponential scaling is that this model only introduces one additional parameter to the default model of hadronization in Herwig 7, and it does not introduce any extra parameters if one splits the fission and decay parameters. Thus we avoid a proliferation of parameters in the new model, while still having a natural mechanism to allow for event-by-event fluctuations in strangeness production.

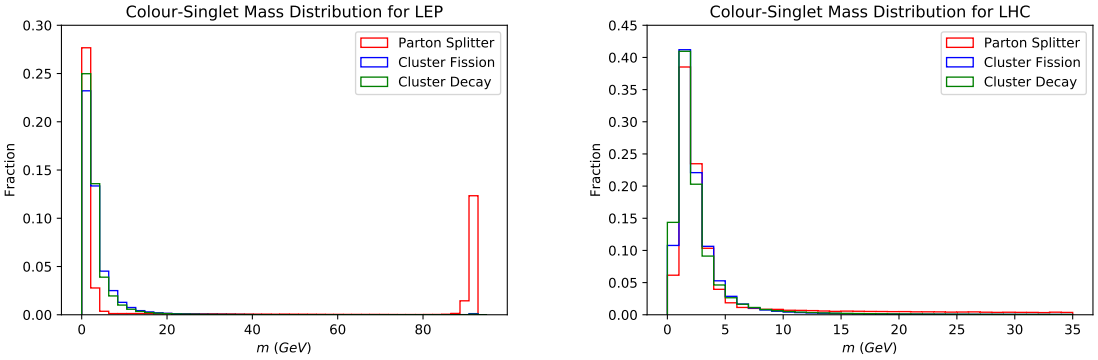

Figure 6.3: Mass distributions for colour-singlet systems immediately before the Parton Splitter, Cluster Fissioner, and Cluster Decayer steps in LEP $\sqrt{s} = 91.2$ GeV and LHC MB events at $\sqrt{s} = 7$ TeV.

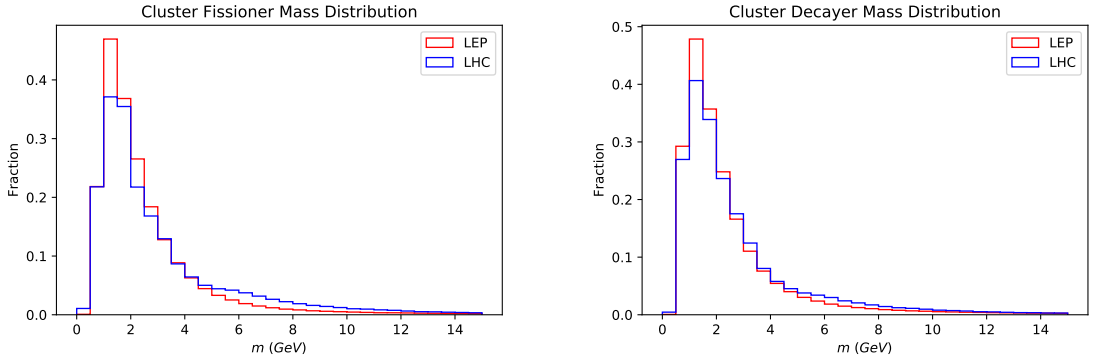

Figure 6.4: Comparison of LEP and LHC MB mass spectra of clusters immediately before cluster fission and cluster decay.

6.3.1 Non-perturbative gluon splitting

At the end of the parton shower, instead of immediately splitting the gluons into $q\bar{q}$ pairs with the species determined by their given weights, we collect the various colour-singlet systems in the event, what we call *pre-clusters*. While colour pre-confinement dictates that the mass distribution of clusters is independent of the hard energy scale,

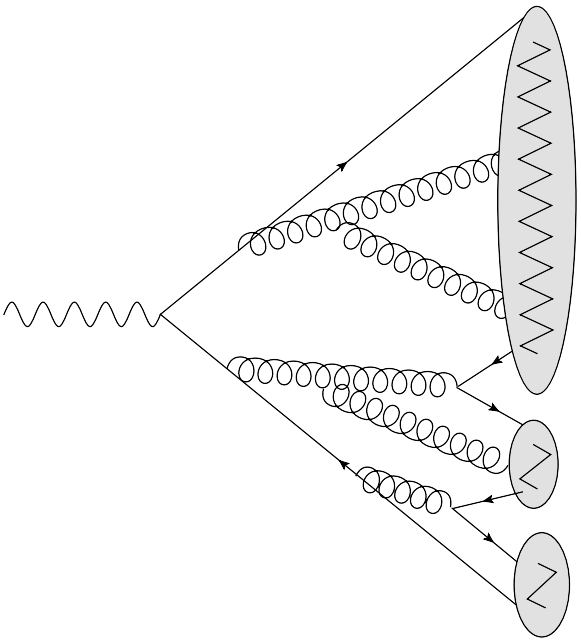

Figure 6.5: Schematic topology of colour-singlets that can occur from perturbative gluon and quark shower splitting, before the gluons undergo non-perturbative splitting.

there are no such constraints on the masses of the colour-singlet pre-clusters. As shown schematically in Fig. 6.5, a parton shower can produce gluons and quark-antiquark pairs at a perturbative level, separating the event into a number of different pre-clusters with a variety of masses.

Every gluon in the same pre-cluster will get the same weight, and thus have the same mass measure for strangeness production. But since the quark species is picked probabilistically, this does not mean that all the gluons will produce strange quark-antiquark pairs. The following constraint from default Herwig 7 still applies: if a single gluon cannot access the phase space necessary to split into a $s\bar{s}$ pair, it will undergo the usual splitting into up or down quarks. The characteristic mass scale for pre-clusters will unfortunately depend on the type of collider one uses. As shown in Fig. 6.3, there is a very broad tail for the LHC due to the different pre-clusters being produced. This is a by-product of the dense and complicated final state environment of high-energy proton-proton collisions. At LEP, there is one peak for the pre-cluster mass distribution close to 91.2 GeV, corresponding to events where gluons are emitted only from the outgoing $q\bar{q}$ legs from the hard scattering process and a proliferation of the distribution when the pre-cluster masses approach zero.

6.3.2 Cluster fission & decay

At the cluster fission and cluster decay stage, the colour-singlet corresponds to the cluster itself. We allow the characteristic mass scale and characteristic production probability to be different for the two hadronization stages. As shown in Fig. 6.4, the typical cluster masses during the cluster fission and cluster decay stage are roughly similar for both LEP and LHC. We note that Fig. 6.3 and Fig. 6.4 are benchmarks of the typical colour-singlet invariant masses, and thereafter plotted without turning on the exponential scaling, which would change the mass distribution slightly.

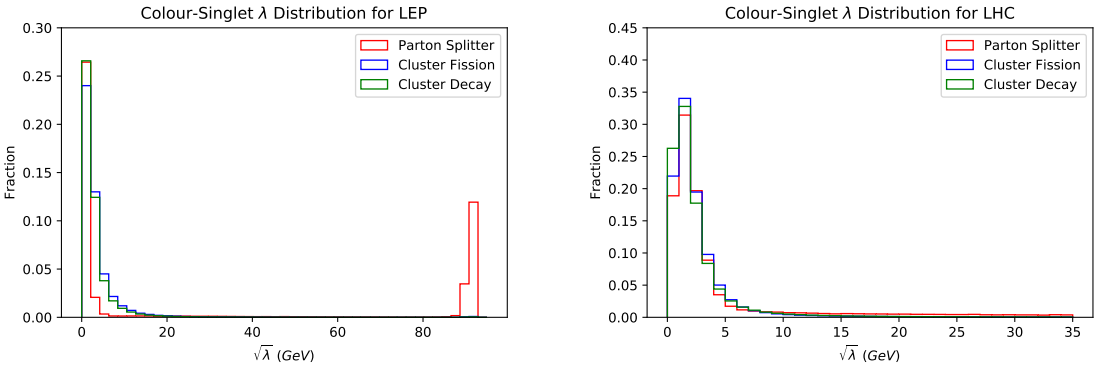

Figure 6.6: Threshold mass, λ, distributions for colour-singlet systems immediately before the Parton Splitter, Cluster Fissioner, and Cluster Decayer steps at LEP events at $\sqrt{s} = 91.2$ GeV and LHC MB events at $\sqrt{s} = 7$ TeV.

6.3.3 Colour-singlet masses

In the previous sections, we have used the invariant mass m^2 of the colour-singlet systems as the mass measure in Eq. (6.1). The invariant mass m^2 neglects the massive nature of the partons in the pre-clusters and clusters. Given two colour-singlets of the same invariant mass, the singlets with lighter constituents should more readily be able produce $s\bar{s}$ pairs from the vacuum.

To remove the biasing effects of massive constituents, we use another mass measure which we call λ measure. We define the λ measure as

$$\lambda = m_{cs}^2 - \left(\sum_i m_i\right)^2, \tag{6.2}$$

where m_{cs}^2 is the invariant mass of the colour-singlet system, and m_i are the masses of the constituents of the colour-singlet. The λ measure replaces the mass-based denominator $f(m^2)$ in Eq. (6.1). We present the distributions of the λ measure for each of the hadronization stages in Fig. 6.6, and a comparison between the distributions of the two mass measures in Fig. 6.8 and Fig. 6.7. The λ measure has the feature that produced $s\bar{s}$ pairs at the gluon splitting level do not propagate as a strangeness enhancement factor further into the hadronization process.

6.4 Analysis

The tunable parameters of the mass-based scaling model are the respective mass-measures m_0 defined by Eq. (6.1) for each stage of the hadronization. We name them `MassScale` (for gluon splitting), `FissionMassScale` (for cluster fissioning), and `DecayMassScale` (for cluster decay). We first tune these 3 parameters to the same identified strange particle yields at LEP and LHC as in Sec. 6.2. The outcome of the tuning procedure for the relevant parameter values is shown in Tab. 6.3 and Tab. 6.4 for LEP and LHC MB events, for both the invariant mass measure and the λ measure.

With the three new characteristic mass scales, we are able to improve the description of all observables considered in the tuning. Especially for LHC observables, as shown

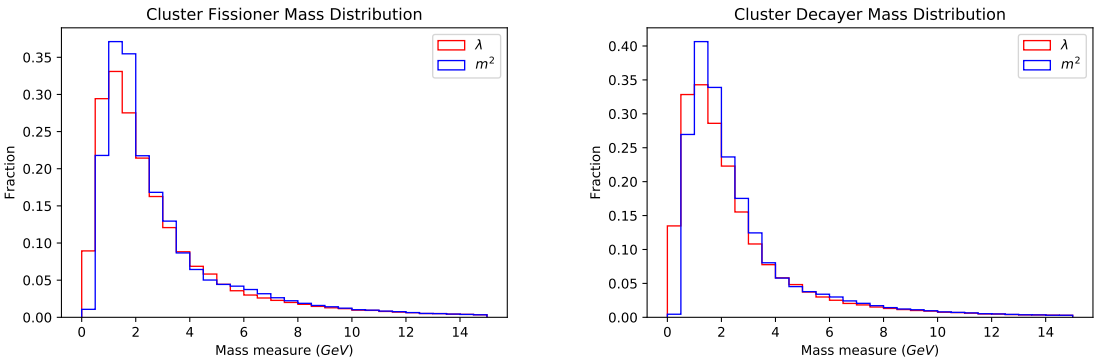

Figure 6.7: Comparison of the two different mass measures for the cluster fission and cluster decayer stages respectively for LEP events at $\sqrt{s} = 91.2$ GeV.

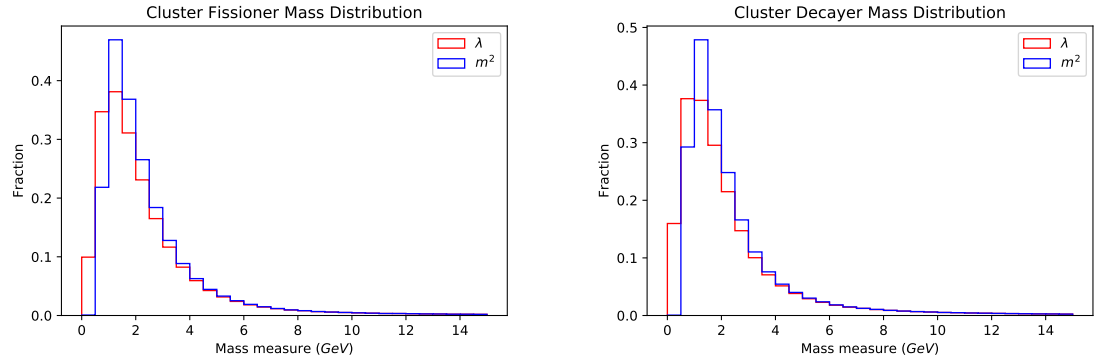

Figure 6.8: Comparison of the two different mass measures for the cluster fission and cluster decayer stages respectively for LHC MB events at $\sqrt{s} = 7$ TeV.

in Fig. 6.9, where we compare the two different mass measures after tuning, with the Monash tune for Pythia.

Although the tuning recommends different values for the usage of the simulation of LHC and LEP events, it is feasible to use the set of parameters obtained from the tuning to LHC data and obtain improved results for LEP observables. This is not possible by using the default model of Herwig 7, which simply uses a fixed probability to produce strange quarks. We show this in Fig. 6.10.

6.4.1 Discussion

The default version of Herwig 7 does not allow for strange quark production during the gluon splitting stage. By allowing this process, improvements can be seen in all of the observables considered.

There are three key differences between the LEP and LHC environments during hadronization. First, LEP has a much lower energy scale than the LHC, naturally limiting the possible distribution of colour-singlet masses at the stage of non-perturbative gluon splittings. As a result, a direct comparison between LEP and LHC in the new model is not straightforward. Second, while LEP and LHC events may have very similar cluster mass distributions, the number of clusters is far higher for the latter.

Invariant Mass	LEP	LHC
Gluon Splitting	97	48
Cluster Fission	3	22
Cluster Decay	23	4

Table 6.3: Results for the tuned characteristic mass scales m_0, in units of GeV, of the new model using the invariant mass of a colour-singlet object for LEP and LHC tunes respectively.

λ Measure	LEP	LHC
Gluon Splitting	72	37
Cluster Fission	4	20
Cluster Decay	16	10

Table 6.4: Results for the tuned characteristic mass scales m_0, in units of GeV, of the new model using the λ measure (defined in Eq. (6.2)) of a colour-singlet object for LEP and LHC tunes respectively.

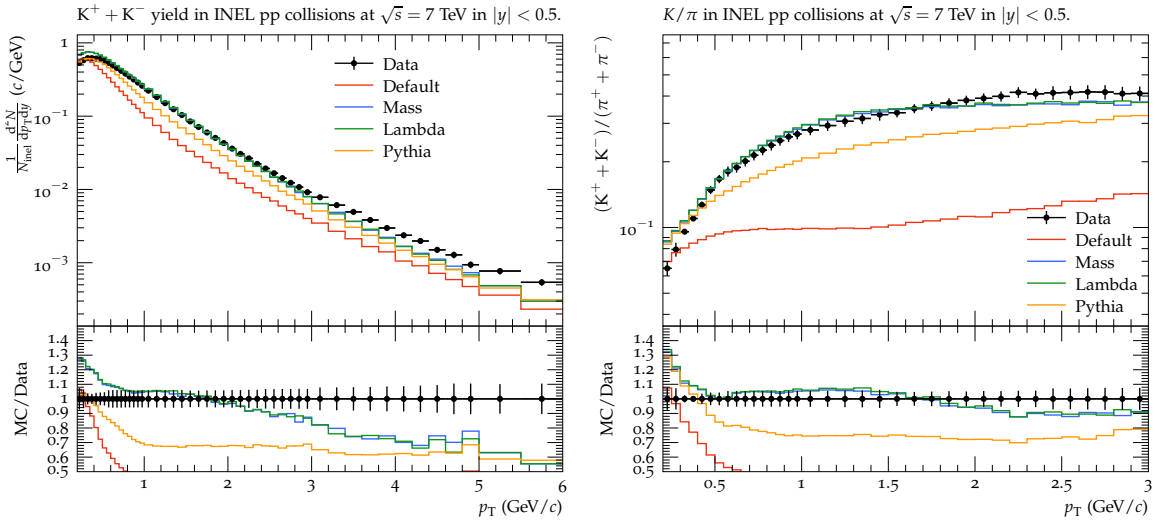

Figure 6.9: $K^+ + K^-$ yield and K/π ratio as measured by ALICE [90] at 7 TeV. Shown is a comparison between the default version of Herwig 7 (without baryonic reconnection), i.e. static production of strangeness, the new approach which introduces dynamical strangeness production with the two different measures (Mass and Lambda) and Pythia with the Monash tune.

Similarly, at the pre-cluster level, LEP events often produce pre-clusters that span the entire final state with a centre-of-mass energy of $\sqrt{s} = 91.2$ GeV as shown in Fig. 6.3. These pre-clusters occur when no perturbative gluon splitting during the parton shower evolution took place. This results in the majority of events either having enhanced strangeness production at the non-perturbative gluon splitting stage, or none at all. Hence a flat weight at this stage in hadronization can only be justified for LEP simulations. Third, and related to the previous two, LEP is a much cleaner environment in terms of hadronic activity. For lepton-lepton collisions, there are neither multiple parton interactions, nor much effect from colour reconnection. However, in proton collisions, these are both vital phases of the simulation that drastically change the mass topology of the event.

We translate the characteristic mass scales from Tab. 6.3 and Tab. 6.4 into an effective expectation value for the weights of the two mass measures. For LEP events, the

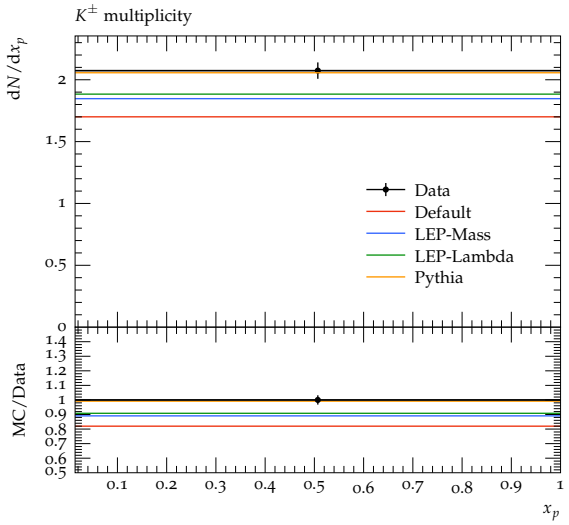

Figure 6.10: Measurement of $K^{\pm}$ multiplicities at SLD [113] $\sqrt{s} = 91.2$ GeV. We show a comparison between the default Herwig 7 model and the dynamical strangeness production where we used the LHC-tuned parameters (see Tabs. 6.3 and 6.4) and Pythia with the Monash tune.

$E(w_s)$ at LEP	Mass	λ
Gluon Splitting	0.096	0.164
Cluster Fission	0.297	0.166
Cluster Decay	0.009	0.016

Table 6.5: Expected value of strangeness production weight of the new model in LEP events at $\sqrt{s} = 91.2$ GeV, comparing the invariant mass results with the λ measure results

$E(w_s)$ at LHC	Mass	λ
Gluon Splitting	0.555	0.571
Cluster Fission	0.018	0.020
Cluster Decay	0.153	0.041

Table 6.6: Expected value of strangeness production weight of the new model in LHC MB events at $\sqrt{s} = 7$ TeV, comparing the invariant mass results with the λ measure results

results are shown in Tab. 6.5. We see that for the invariant mass measure, cluster fissioning has the highest expectation value, while the expectation value during the cluster decay stage is suppressed. For the λ measure, the expectation values during gluon splitting and cluster fission are roughly the same and strangeness production during cluster decay is also suppressed. It should be noted that aside from the gluon splitting weights, there is no direct translation between the kinematic picture and the default model. The expectation values, however, give an idea of the average weights for strangeness production. The weight for strangeness production during gluon splitting at LEP varies between 0 and the maximal value since the pre-clusters are predominantly situated around the two peaks, as shown in Fig. 6.3. For LHC MB events, the expected value for the weights are shown in Tab. 6.6. For both mass measures the gluon splitting stage dominates the strangeness production compared to the two other stages. The suppression of strangeness production during the cluster fission and cluster decay stage compared to the gluon splitting is a hint that colour reconnection plays a non-trivial role

in producing strange hadrons. Our new kinematic model uses a mass-based scaling, but colour reconnection aims to lower the cluster masses to some local minimum, meaning that it is in direct conflict with our considerations. The results in Tabs. 6.5 and 6.6 suggest that the λ measure is able to bridge the divide between the two types of collision at LHC and LEP better than the invariant mass measure.

We compared the results of the new model with Pythia and the Monash tune in Fig. 6.9 and Fig. 6.10. While the Monash tune aims to describe a number of observables other than the strangeness production rate in Pythia, it is tuned to both LEP and LHC data [110], making it an apt benchmark for this discussion. We observe that the new model performs marginally better than Pythia, and significantly better than the default version of Herwig 7, when trying to describe the $K^{\pm}$ and drastically better on both counts for the K/π ratio yields, as shown in Fig. 6.9. However, in the low-$p_{\perp}$ region, both Pythia and the new model overestimate the data. When using LHC MB tuned parameters for LEP simulations, the new model outperforms the default Herwig 7 model, but Pythia is able to describes the data better, as shown in Fig. 6.10.

While we have not fully solved the discrepancy between the weights for LEP and LHC strangeness production, we have achieved two results: first, we have narrowed the gap between the weights of the two types of collision, and in particular, the new model can be used with parameter values tuned to LHC MB observables to better describe LEP data. Second, we have made the first steps towards a more sophisticated treatment of strangeness production during hadronization at the low-energy scale in Herwig 7.

6.5 Conclusion

We have introduced a three-part model that scales the probability for strangeness production during the hadronization phase of the event generation in Herwig 7. The scaling is directly controlled by the mass of the corresponding colour-singlet system at each step of the hadronization. With this mechanism, we allow for greater fluctuations in the production of strange pairs on an event-by-event basis.

We have studied the mechanism for non-perturbative strangeness production in detail and found that the current fixed probability model is irreconcilable with both LEP and LHC data. After allowing a mass-based scaling, and tuning the parameters to LEP and LHC data, we find that we are able to narrow the gap between the two collider types. We also provide expectation values for strangeness production, which capture the average values for event-by-event fluctuations. We have compared the new model with the Pythia Monash tune to strangeness observables at both LEP and LHC and find that our model improves the description of all observables considered.

We note that we have not considered the production of strange baryons, the production of which has been shown to be increased by creating baryonic clusters at the colour reconnection stage (see Ch. 4). Baryonic clusters, which are heavier by nature, would modify the new model's strangeness production rates. Understanding the interplay between the new model and colour reconnection will be left for future work

The model for kinematic strangeness production was released with Herwig 7, version 7.2 [94].

CHAPTER 7

Improving the Modelling of Multiple Parton Interactions in Herwig 7

Multiple parton interactions and diffractive processes are responsible for the majority of the possible collision types during hadronic collisions. The modelling of such processes is therefore crucial in the description of MB and UE data. The general properties of MB and UE measurements can often be described satisfactorily, but as measurements become more and more differential, current models are quickly reaching the limit of capabilities. In this section, we highlight several changes to the multiple parton interaction model of Herwig 7 which are necessary for a more physically sound framework of multiple parton interactions and diffraction.

In Sec. 7.1 we improve the sampling of the kinematics of the ladder partons from the soft interactions presented in Sec. 3.5. In Sec. 7.2 we include the model for diffraction into the eikonal model in a more sensible way and in Sec. 7.3 we tune the revised model to data available from UE and MB measurements covering an energy range between $\sqrt{s} = 200$ GeV and $\sqrt{s} = 13$ GeV. We conclude our study with a quick discussion in Sec. 7.4.

7.1 Soft Interactions

The kinematics of the ladder partons are sampled according to the algorithm presented in Ref. [53]. Although the model leads to a significant improvement in the description of UE and MB data, it has two physical problems. First, the algorithm that distributes the ladder partons evenly in rapidity calculates a longitudinal momentum fraction which is given to each parton in the ladder. By starting with both remnants at the same time, the ladder partons are always produced as pairs where both partons on opposing sides are given the same longitudinal momentum fraction[5]. This procedure leads to highly

[5]For a review of the algorithm, see Ref. [53].

anti-correlated mini-jet events as was pointed out in Ref. [115]. Second, the transverse momenta of all ladder partons are sampled from the distribution defined by Eq. (3.28). This creates a bias towards larger $p_\perp$ values for the hardest parton in the ladder and the shape of the calculated $\mathrm{d}\sigma/\mathrm{d}p_\perp$ distribution is not reproduced. This also results in too many particles originating from the ladder with a $p_\perp$ shifted towards larger values. To address these issues, we change the algorithm responsible for the kinematics of the ladder partons.

Instead of sampling the longitudinal momentum fraction given to each ladder parton, we first distribute the partons flat in rapidity and then calculate the longitudinal momentum fraction given to each parton, such that momentum and energy are conserved, based on a modified version of the Jadach algorithm [105]. This resolves the occuring anti-correlations in soft mini-jet events. We further change the sampling of the transverse momenta of the ladder partons and require that the transverse momenta of the hardest parton in the ladders should reproduce the calculated $\mathrm{d}\sigma/\mathrm{d}p_\perp$ distribution from Eq. (3.28).

For the new kinematics of the ladder partons, we use the modified version of the Jadach algorithm which was used in the old UA5 model of Herwig [3,116]. Instead of generating clusters similar to the UA5 model, we stick to the microscopic description on the parton level. This has shown to be more physical and effective than a simple parametrization of the final state used in the UA5 model.

In the following we describe each step of the new algorithm for the generation of the kinematics of the ladder partons. The starting point of the algorithm is the number of soft interactions N_{soft} calculated within the eikonal model (see Sec. 3.5). For every soft interaction we generate a ladder of partons where the available energy for the ladder partons depends on the centre-of-mass energy E_{CM} of the two remnants

$$E_{\mathrm{CM}} = \sqrt{(p_{r1} + p_{r2})^2}, \tag{7.1}$$

where p_{r1} and p_{r2} are the four-momenta of the remnants. To simulate the soft interactions, we do the following steps N_{soft} times.

1. We sample the number of ladder partons N from a Poissonian distribution with mean value given by Eq. (3.25).

2. The energy fraction available for the N ladder partons is calculated according to

$$E_0 = xE_{\mathrm{CM}}, \tag{7.2}$$

   where x is sampled from a Gaussian with mean

$$x = e^{\frac{\Delta Y}{N+1}}. \tag{7.3}$$

   If the remaining remnant energy is not sufficient to sustain at least the rest-mass of the remnants, $E_{\mathrm{CM}} - E_0 < m_{r1} + m_{r2}$, we sample another energy fraction x. If this fails 10 times, we assume that the remaining energy in the remnants is not sufficient to generate additional soft interactions and we terminate the algorithm.

3. We generate the masses $m_{1,\ldots,N}$ of the N ladder partons where 2 partons are given quark masses and $N-2$ partons are given the non-perturbative gluon mass. It is essential to store the parton masses since the modified Jadach algorithm deals with massless partons.

4. The available energy to generate the momenta of the ladder partons is then given by

$$E_0' = E_0 - \sum_{i=1}^{N} m_i, \tag{7.4}$$

where we require that $E_0' \overset{!}{>} 0$.

5. In the next step we generate the transverse momenta $p_{\perp,1,...N}$ of the ladder partons. To make sure the transverse momenta are not biased towards higher values we sample the first $p_\perp$ from Eq. (3.28)

$$\frac{\mathrm{d}\sigma_{\mathrm{soft}}}{\mathrm{d}p_\perp} \sim p_\perp e^{-\beta(p_\perp^2 - p_{\perp,\mathrm{min}}^2)}, \tag{7.5}$$

and the transverse momenta of the partons $2, ..., N-1$ from a flat distribution below $p_{\perp,1}$. This is necessary to reproduce the shape of the $\mathrm{d}\sigma_{\mathrm{soft}}/\mathrm{d}p_\perp$ distribution for the hardest parton in the ladder. The transverse momenta are translated to p_x and p_y values by randomly sampling the azimuthal angle between 0 and 2π. To account for momentum conservation in the subspaces

$$\sum_{i=1}^{N} p_{i,x} = 0, \qquad \sum_{i=1}^{N} p_{i,y} = 0, \tag{7.6}$$

we choose the transverse momentum $p_{\perp,N}$ such that

$$p_{N,x} = -\sum_{i=1}^{N-1} p_{x,i}, \tag{7.7}$$

$$p_{N,y} = -\sum_{i=1}^{N-1} p_{y,i}, \tag{7.8}$$

where we furthermore ensure, that $p_{\perp,N} \overset{!}{<} p_{\perp,1}$.

6. We shuffle the order of partons to remove the bias we introduce by calculating $p_{\perp,N}$ from Eq. (7.7) and Eq. (7.8).

7. Now that all transverse momenta are produced we generate the rapidities of the partons according to the procedure outlined in Ref. [105]. We first generate N random numbers ξ_i in the range $0 \leq \xi_i \leq 1$. The individual rapidities y_i are then obtained from the ξ_i with a linear transformation

$$y_i = Z + \xi_i Y, \tag{7.9}$$

where Z and Y are unknown variables. We can use the following constraints to determine the values of Z and Y. In the centre-of-mass frame, the energy E_0' can be written in the massless parton case by using $y_i = \ln[(E_i + p_{z,i})/p_{\perp,i}]$ and energy-momentum conservation as

$$E_0' = e^Z \sum_{i=1}^{N} p_{\perp,i}\, e^{\xi_i Y}, \tag{7.10}$$

and the squared energy as

$$E_0'^2 = \left(\sum_{i=1}^{N} p_{\perp,i} e^{\xi_i Y}\right)\left(\sum_{i=1}^{N} p_{\perp,i} e^{-\xi_i Y}\right). \tag{7.11}$$

We determine Y from Eq. (7.11) by solving

$$\ln(E_0'^2) = \ln\left[\left(\sum_{i=1}^{N} p_{\perp,i} e^{\xi_i Y}\right)\left(\sum_{i=1}^{N} p_{\perp,i} e^{-\xi_i Y}\right)\right], \tag{7.12}$$

and then calculate Z from Eq. (7.10). Once Y and Z are determined, we can construct the longitudinal momenta $p_{z,1,\ldots,N}$ of the ladder partons.

8. Next, we sort the ladder partons based on their rapidity. The parton with the largest rapidity has to be a quark (antiquark) and the parton with the smallest rapidity has to be an antiquark (quark). We chose the flavour of the quarks between up and down quarks with unit weight. It is necessary to have a quark-antiquark pair with the maximum and minimum rapidity in the ladder to be able to colour connect partons adjacent in rapidity. We give the massless partons their masses based on their type and re-scale the energy of the partons such that energy-momentum conservation is not violated.

9. We boost the momenta of the ladder partons into the centre-of-mass frame of the two remnants and re-scale the four-momenta of the two remnants such that

$$p_{r1} + p_{r2} = \sum_{i=1}^{N} p_i + p'_{r1} + p'_{r2}, \tag{7.13}$$

where p'_{r1} and p'_{r2} are the re-scaled remnant four-momenta. As a cross-check we require that

$$(p_{r1} + p_{r2})^2 - \left(\sum_{i=1}^{N} p_i + p'_{r1} + p'_{r2}\right)^2 \overset{!}{<} 10^{-10}. \tag{7.14}$$

10. Next, we need to account for the colour connections between the ladder partons. The remnants always carry anticolour and are either colour connected to the hard part of the event or colour connected to a quark which was extracted from the hadron to initiate a collision with no main hard process present. We set the colour connections of the ladder partons such that adjacent partons are colour connected with each other. After the colour connections are set, we continue with the next soft interaction until there is no energy in the remnants left, or until the requested number of soft interactions N_{soft} is reached.

The partons originating from the ladder are no longer highly anti-correlated since we explicitly distribute the partons flat in rapidity and then proceed to construct the longitudinal momenta. In Fig. 7.1 we show the observable from Ref. [115], where we compare the old model with the improved model. We see that with the updated algorithm, the unphysical anti-correlations in the region $\eta \sim 2$ disappear. The changes to the $p_\perp$ sampling has a significant effect on the low multiplicity region of the $\langle p_\perp \rangle$ vs. N_{ch} observable as shown in Fig. 7.2. While the improved algorithm leads to an

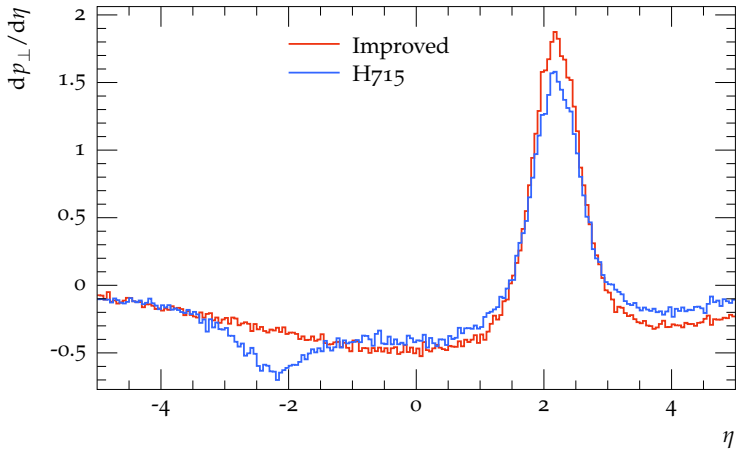

Figure 7.1: Plot of the unphysical mini-jet correlations from Ref. [115]. Shown is a comparison between the model for multiple parton interactions from Herwig 7.15 and the improved version of the model.

improvement of the description in the $N_{\mathrm{ch}} < 40$ region, the description seems to worsen for very high multiplicities. However, the description of this region has to be taken with a grain of salt. From charged multiplicity distributions, we know that events with very high multiplicities $N_{\mathrm{ch}} > 140$ are suppressed by a factor of at least 10^{-6} (see e.g Ref. [67]). Collecting significant statistics in these regions poses an additional challenge. Furthermore, this region also signifies the onset of a regime where the usual employed tactic to layer multiple interactions on top of each other seems to break down. This is also due to the fact that the modelling of high-multiplicity events in proton-proton collisions is not yet completely understood.

7.2 Diffraction

In Herwig 7 the models for multiple parton interactions and diffraction are modelled independently. However, they are physically connected via the cross-sections calculated in the eikonal model. To connect the diffractive model to the model for multiple parton interactions, we have to ensure that the inelastic cross-section σ_{inel} from Eq. (3.17) only sums up to a cross-section where the diffractive cross-section has been subtracted

$$\sigma_{\mathrm{inel}} \equiv \sigma_{\mathrm{inel-nondiff}} = \sigma_{\mathrm{tot}} - \sigma_{\mathrm{el}} - \sigma_{\mathrm{diff}}. \tag{7.15}$$

The diffractive cross-section is given by a fraction R_{diff} of the total cross-section

$$\sigma_{\mathrm{diff}} = R_{\mathrm{diff}}\sigma_{\mathrm{tot}}. \tag{7.16}$$

With the fraction of diffractive events, we can adjust the cross-sections in the multiple parton interaction model such that the sum after eikonalization gives the correct total cross-section [117]. This parameter also gives an excellent handle on the fraction of diffractive events and can be tuned to observables sensitive to diffraction like the differential rapidity gap cross-section in the forward region $\mathrm{d}\sigma/\mathrm{d}\Delta\eta^F$ as measured by ATLAS at $\sqrt{s} = 7$ TeV. In Fig. 7.3 we show the $\mathrm{d}\sigma/\mathrm{d}\Delta\eta^F$ observable for three different values of the R_{diff} parameter, which is called DiffractionRatio in the plot. We see

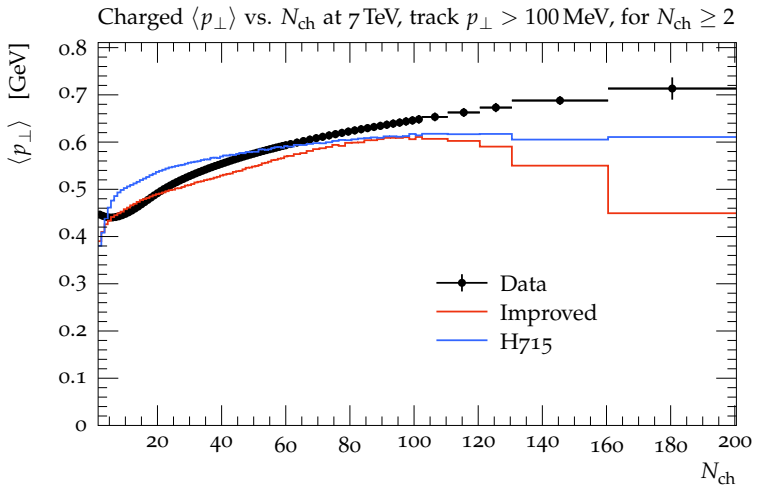

Figure 7.2: Comparison between the old and the improved version of the model. The plot shows the average transverse momentum $\langle p_\perp \rangle$ against the charged multiplicity N_{ch} as measured by ATLAS [67].

that varying R_{diff} directly modifies the differential cross-section with respect to events containing large rapidity gaps $\Delta\eta^F$, but leaves the region $\Delta\eta^F < 2$ invariant. This region is dominated by non-diffractive particle production on which the parameter has no influence.

7.3 Tuning

Since there were several changes to the underlying structure of the multiple parton interaction model and the soft part of the model itself, a re-tune of the model parameters is necessary. We tune the model to available MB and UE data covering the centre-of-mass energies between $\sqrt{s} = 200$ GeV and $\sqrt{s} = 13$ TeV [56, 67, 67, 75, 89, 89, 90, 119, 120, 120, 121, 121, 122, 122–130]. For the tuning we consider the following parameters, which determine the properties of the multiple parton interaction model:

- the minimum transverse momentum $p_\perp^{min}$,
- the inverse proton radius squared μ^2,
- the soft ladder parameter N_0,
- the fraction of the diffractive cross-section from the total cross-section R_{diff},
- the two colour reconnection probabilities p_{Reco} and $p_{RecoBaryonic}$.

The energy dependence of the $p_\perp^{min}$ parameter is parametrized according to the following power law [72]

$$p_\perp^{min}(s) = p_{\perp,0}^{min} \left(\frac{\sqrt{s}}{E_0} \right)^b . \tag{7.17}$$

The parameters of the $p_\perp^{min}$ parametrization were tuned to $\sqrt{s} = 900$ GeV and $\sqrt{s} = 7$ TeV in Ref. [53]. The parametrization Eq. (7.17) results in a hard cross-section which exceeds the total cross-section for small centre-of-mass energies. This in return

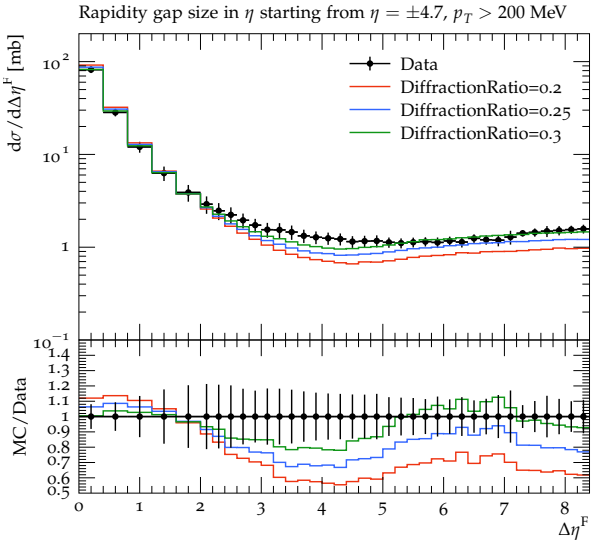

Figure 7.3: Comparison of the differential cross-section measured in Ref. [118] to different choices of the DiffractionRatio parameter.

$\sqrt{s}$ [GeV]	200	500	900	1800	7000	13000
$p_\perp^{\min}$ [GeV]	1.80	1.6	1.75	1,95	2.91	3.5
μ^2 [1/GeV^2]	1.0-2.2	1.1-2.5	1.0-2.0	1.0-1.6	1.05-1.15	1.0-1.2

Table 7.1: Tune values for $p_\perp^{\min}$ and the ranges of μ^2.

terminates the multiple parton interaction model since a negative soft cross-section would be required to restore unitarity.

To make certain that the model works at all centre-of-mass energies, we start the tuning with a reassessment of the energy dependence of the $p_\perp^{\min}$ parameter in combination with the μ^2 parameter. These two parameters are the main parameters of the model for multiple parton interactions and are responsible for the interplay between hard and soft interactions by directly stirring the relevant cross-sections. We tune both values to MB and UE data at $\sqrt{s} = 200$ GeV [119, 120], $\sqrt{s} = 500$ GeV [120], $\sqrt{s} = 900$ GeV [67,89,121,122], $\sqrt{s} = 1.8$ TeV [123–127], $\sqrt{s} = 7$ TeV [56,67,75,89,90,121,122] and $\sqrt{s} = 13$ TeV [128–130]. For each centre-of-mass energy we generate 200 runs. Every run consists of 100,000 events with randomly selected parameter values in the ranges $p_\perp^{\min} \in (1.5, 4.0$ GeV) and $\mu^2 \in (0.8, 3.0$ GeV$^{-2})$. From these 200 runs a subset of 150 runs is used 100 times for the tuning. For each of these run combinations, the χ^2/N_{dof} value is calculated. While the μ^2 values are spread within the given parameter ranges for lower collision energies, the $p_\perp^{\min}$ parameter favours distinctive values for each centre-of-mass energy.

In a similar approach as in Ref. [72], we fix the μ^2 parameter to the tightly constrained parameter values for $\sqrt{s} = 7$ TeV and $\sqrt{s} = 13$ TeV. With μ^2 fixed, we then re-tune $p_\perp^{\min}$ to the different centre-of-masss energies. The $p_\perp^{\min}$ value which results in the best description of the data and the spread of the μ^2 parameter resulting from the different run combinations are shown in Tab. 7.1.

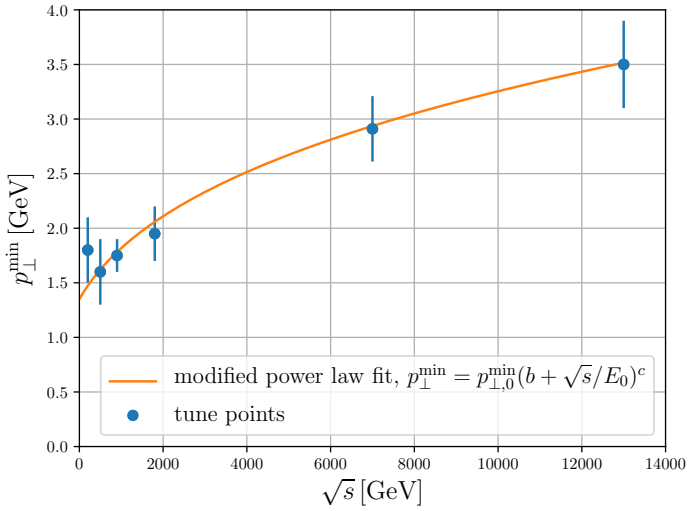

Figure 7.4: Energy parametrization of the parameter $p_\perp^{\min}$. The blue dots represent the optimal value. The error bars represent the spread of the $p_\perp^{\min}$ for different run combinations.

$p_{\perp,0}^{\min}$ [GeV]	b [GeV]	c	E_0 [GeV]
2.87	622	0.31	7000

Table 7.2: Parameter values used in the modified energy parametrization of $p_\perp^{\min}$ (see Eq. (7.18)).

In order to achieve a good description of MB and UE data over the considered energy range between $\sqrt{s} = 200$ GeV and $\sqrt{s} = 13$ TeV, we modify the existing power law parametrization of the $p_\perp^{\min}$ parameter and introduce an energy offset to account for $p_\perp^{\min}$ values at $\sqrt{s} < 900\,\text{GeV}$. The modified power law reads

$$p_\perp^{\min}(s) = p_{\perp,0}^{\min}\left(\frac{b+\sqrt{s}}{E_0}\right)^c, \tag{7.18}$$

where $p_{\perp,0}^{\min}$, b and c are parameters which we need to determine. In Fig. 7.4, we show the tuned $p_\perp^{\min}$ values at their respective centre-of-mass energy and the resulting fit of the modified power law. Since we opt for an improved description at higher centre-of-mass energies, we demand that the parametrization must reproduce the $p_\perp^{\min}$ values for $\sqrt{s} = 7\,\text{GeV}$ and $\sqrt{s} = 13\,\text{TeV}$. The fitted values of the parametrization Eq. (7.18) are listed in Tab. 7.2.

With the modified $p_\perp^{\min}$ parametrization, the model for multiple parton interactions works at all centre-of-mass energies considered and the reoccurring issue that the hard cross-section exceeds the total cross-section at small $\sqrt{s}$ is no longer present.

We tune the remaining parameters N_0, R_{diff}, p_{reco}, and $p_{\text{recoBaryonic}}$ to MB and UE data at $\sqrt{s} = 7\,\text{TeV}$ and $\sqrt{s} = 13\,\text{TeV}$. We refrain from tuning to single observables and instead focus on an overall good description of the data. The resulting set of parameters, which were tuned with the new parametrization for $p_\perp^{\min}$ are listed in Tab. 7.3. In

addition to the energy-independent tune, we provide individual tunes for the full set of parameters for $\sqrt{s} = 7\,\text{TeV}$ and $\sqrt{s} = 13\,\text{TeV}$. The resulting parameter values are listed in Tab. 7.3. The value of the diffractive ratio is mainly driven by the $\mathrm{d}\sigma/\mathrm{d}\Delta\eta^F$ observables from Ref. [56,131] and the diffractive cross-sections measurements, σ_{SD}, σ_{DD} from Ref. [65] at $\sqrt{s} = 7$ TeV. More data of diffractive events at different centre-of-mass energies is needed to correctly assess the energy dependence of the diffractive cross-section (see e.g. Ref. [132]). The probability for baryonic reconnection is mainly driven by flavour observables at $\sqrt{s} = 7$ TeV. As can be seen in Tab. 7.4, the reconnection probability $p_{\text{RecoBaryonic}}$ drops from $\sqrt{s} = 7$ TeV to $\sqrt{s} = 13$ TeV due to the absence of MB measurements with repsect to flavour observables at $\sqrt{s} = 13\,\text{TeV}$.

Parameter	Value
N_0	0.6838
R_{Diff}	0.187
p_{Reco}	0.970
$p_{\text{RecoBaryonic}}$	0.626
μ^2	1.1
$p_{\perp,0}^{\text{min}}$	2.82
b	622.203
c	0.31
E_0	7000

Table 7.3: Parameters of the energy independent UE and MB tune with the parametrization of $p_\perp^{\text{min}}$ from Eq. (7.18).

Parameter	7 TeV	13 TeV
N_0	0.601	0.699
R_{Diff}	0.2574	0.167
p_{Reco}	0.99	0.811
$p_{\text{RecoBaryonic}}$	0.898	0.496
μ^2	1.1	1.1
$p_\perp^{\text{min}}$	2.827	3.97

Table 7.4: Energy dependent parameter values of the UE and MB tune for $\sqrt{s} = 7$ TeV and $\sqrt{s} = 13$ TeV.

7.4 Conclusion

We introduced several changes to the model for multiple parton interactions in Herwig 7. We replaced the sampling algorithm of the ladder partons with an algorithm that directly distributes the partons flat in rapidity and then calculates the momenta of the partons. Additionally, we changed the sampling of the transverse momenta of the ladder partons such that the $\mathrm{d}\sigma/\mathrm{d}p_\perp$ distribution is reproduced for the hardest parton in the ladder. These changes dispose of the observed anti-correlations and also improve the description of data. We also took one step closer to the integration of the diffractive model in Herwig 7 by directly linking it with the cross-sections calculated in the eikonal model. We tuned the resulting model to available MB and UE data covering an energy range between $\sqrt{s} = 200$ Gev and $\sqrt{s} = 13$ TeV. To account for a good description at all centre-of-mass energies, we change the existing parametrization of the $p_\perp^{\min}$ parameter. The new energy parametrization results in a good description of all observables available with only one set of parameters. Albeit due to the nature of the combined tune, there are aspects which are not described satisfactorily. Dedicated single observable tunes will of course always outperform combined tunes.

The changes to the model for multiple parton interactions presented in this section and the tuned parameter values including the modified energy parametrization of the $p_\perp^{\min}$ parameter are set as the default in the new version of Herwig 7.2 [94]. The revised model can be used as a well-motivated baseline describing general properties of MB and UE data over a wide range of centre-of-mass energies.

CHAPTER 8

Summary and Conclusions

In this thesis, we have addressed and improved the simulation of non-perturbative models in the Monte Carlo event generator Herwig 7. Furthermore, we have studied colour reconnection with an approach from first principles and gained insight into the underlying structure of the colour reconnecting effects.

We have implemented a new model for colour reconnection which is based on a geometrical picture in Ch. 4. By allowing the formation of baryonic clusters in combination with strangeness production during non-perturbative gluon splitting, we improve the description of MB data for flavour observables significantly. The model also enables us to describe the enhancement of (multi-)strange hadrons as seen by ALICE in high-multiplicity pp collisions.

In Ch. 5, we approached the topic of colour reconnection from a theoretical point of view by considering soft gluon exchanges and study the resulting cluster configurations in various phase-space regions. We found strong support for colour reconnection models based on geometrical considerations. The theoretical insight we gained can serve as input for the development of more sophisticated colour reconnection models.

We studied the production of strangeness in the context of the cluster hadronization model in Ch. 6. We found that the current model was irreconcilable with both LEP and LHC data. By implementing a three-part model, where the strangeness production rate at each step of hadronization depends on the immediate environment, we allow for more dynamic strangeness production on an event-by-event basis. The new model significantly improves the description of strangeness observables and narrows the gap between the favoured parameter values between LEP and LHC.

In Ch. 7, we introduced several changes to the model for multiple parton interactions of Herwig 7 to account for a more physically sound model. We improved the kinematic sampling of the ladder partons and included the model for diffraction into the eikonal model in a more sensible way. We tuned the model parameters to MB and UE data

covering an energy range from $\sqrt{s} = 200$ GeV to $\sqrt{s} = 13$ TeV. With the changes to the model for multiple parton interactions, we find that we are able to describe all relevant data with one set of parameters.

The new models presented in this thesis have been released with the new version of Herwig 7.2 [94] and prepare Herwig 7 for future studies related to non-perturbative physics.

There are nonetheless many avenues for possible future projects. Since the study of high-multiplicity events in light of the recent ALICE data gained some increasing interest it would be benificial to study the high-multiplicity topologies in greater detail and investigate the correlations observed and compare them to different simulations and models. A possible starting point would be to investigate the differences between the Lund string model, the cluster hadronization model and the Dipsy rope model, which seem to describe the rising trend of (multi)-strange hadrons, and investigate observables which manage to capture differences in the different approaches. What would be also interesting is an investigation of the differences in strangeness production between the cluster hadronization model and other hadronization models.

Furthermore, a systematic tuning of the complete hadronization model to LHC data would also be benificial. On the one hand it may improve the description of hadronic observables even further and on the other hand it would allow us to better restrain the free parameters of the hadronization model, which could lead to conflict with the description of LEP data. A discrepancy in the description of LEP and LHC data could mean that we may perhaps need to distance ourselves from the approach to have a single hadronization model for the simulation of both LEP and LHC. This could signify the breakdown of jet universality and furthermore lead to a reevaluation of factorization approaches with respect to hadronization.

In light of future Herwig 7 developments, it would make sense to include a model for multiple parton interactions where the number of additional interactions directly depends on the sampled impact parameter. The transverse momenta of the additional interactions could then also depend on the multiplicity which should lead to a shift towards larger transverse momenta for high-multiplicity events with many soft and hard interactions and could improve the description in the high-multiplicity tail of the $\langle p_\perp \rangle$ vs. N_{ch} observable. First steps in this direction were undertaken in Ref. [69]. Further down the road is the study of heavy-ion collisions with Herwig 7. Herwig 7 is already able to simulate heavy-ion collision by stacking events on top of each other [133]. However, this approach does not allow for final state correlations which are necessary to explain a multitude of effects seen in heavy-ion collisions [78–83].

Since particles from all possible processes can contribute to a given final state, it is imperative that searches for *new physics* can rely on an excellent modelling of the hadronic background. New approaches and ideas are therefore continuously needed to improve the physics capabilities of Monte Carlo event generators and allow them to provide reliable results in regimes that are dominated by non-perturbative effects.

Acknowledgements (Danksagungen)

Ich möchte mich zunächst bei meinem Doktorvater, PD Dr. Stefan Gieseke, bedanken der mir durch seine geduldige Art in den vergangenen Jahren sehr geholfen hat und immer mit Rat und Tat zur Seite stand.

Des weiteren bedanke ich mich herzlich bei Prof. Dr. Dieter Zeppenfeld für die Übernahme des Korreferats und die Diskussionen während den Institutsseminaren.

Einen ganz besonderen Dank möchte ich an meine Eltern aussprechen die mir die Möglichkeit gegeben haben Physik zu studieren und mich in all meinen Entscheidungen immer unterstützt haben.

I would like to thank Simon Plätzer, Johannes Bellm, Cody Duncan und Andrzej Siódmok for the fruitful collaboration in the past years. It was always a pleasure to work with you and I hope we can continue our collaboration in the future.

Mein besonderer Dank möchte ich nochmal an Simon aussprechen, dafür dass er mir die Möglichkeit gegeben hat eine längere Zeit in Wien zu arbeiten.

Furthrmore, I would like to thank all members of the Herwig and MCnet collaboration for the engaging discussions, interesting meetings and fun workshops.

Ich bedanke mich ausserdem herzlich bei Cody Duncan und Konstantin Asteriadis für das Korrekturlesen der Doktorarbeit.

I thank all members of the ITP for the pleasant time during the past 4 years.

References

[1] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, "An Introduction to PYTHIA 8.2," *Comput. Phys. Commun.* **191** (2015) 159–177, arXiv:1410.3012 [hep-ph].

[2] T. Gleisberg, S. Höche, F. Krauss, M. Schönherr, S. Schumann, F. Siegert, and J. Winter, "Event generation with SHERPA 1.1," *JHEP* **02** (2009) 007, arXiv:0811.4622 [hep-ph].

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

[4] J. Bellm *et al.*, "Herwig 7.0/Herwig++ 3.0 release note," *Eur. Phys. J.* **C76** no. 4, (2016) 196, arXiv:1512.01178 [hep-ph].

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

[6] B. R. Webber, "A QCD Model for Jet Fragmentation Including Soft Gluon Interference," *Nucl. Phys.* **B238** (1984) 492–528.

[7] S. Gieseke, P. Kirchgaeßer, and S. Plätzer, "Baryon production from cluster hadronisation," *Eur. Phys. J.* **C78** no. 2, (2018) 99, arXiv:1710.10906 [hep-ph].

[8] S. Gieseke, P. Kirchgaeßer, S. Plätzer, and A. Siódmok, "Colour Reconnection from Soft Gluon Evolution," *JHEP* **11** (2018) 149, arXiv:1808.06770 [hep-ph].

[9] C. B. Duncan and P. Kirchgaeßer, "Kinematic strangeness production in cluster hadronization," *Eur. Phys. J.* **C79** no. 1, (2019) 61, arXiv:1811.10336 [hep-ph].

[10] H. Fritzsch, M. Gell-Mann, and H. Leutwyler, "Advantages of the color octet gluon picture," *Physics Letters B* **47** no. 4, (1973) 365 – 368. http://www.sciencedirect.com/science/article/pii/0370269373906254.

[11] D. J. Gross and F. Wilczek, "Ultraviolet behavior of non-abelian gauge theories," *Phys. Rev. Lett.* **30** (Jun, 1973) 1343–1346. https://link.aps.org/doi/10.1103/PhysRevLett.30.1343.

[12] H. D. Politzer, "Reliable perturbative results for strong interactions?," *Phys. Rev. Lett.* **30** (Jun, 1973) 1346–1349. https://link.aps.org/doi/10.1103/PhysRevLett.30.1346.

[13] J. C. Collins, D. E. Soper, and G. F. Sterman, "Soft Gluons and Factorization," *Nucl. Phys.* **B308** (1988) 833–856.

[14] J. C. Collins, D. E. Soper, and G. F. Sterman, "Factorization of Hard Processes in QCD," *Adv. Ser. Direct. High Energy Phys.* **5** (1989) 1–91, arXiv:hep-ph/0409313 [hep-ph].

[15] J. C. Collins and A. Freund, "Proof of factorization for deeply virtual Compton scattering in QCD," *Phys. Rev.* **D59** (1999) 074009, arXiv:hep-ph/9801262 [hep-ph].

[16] G. 't Hooft, "A planar diagram theory for strong interactions," *Nuclear Physics B* **72** no. 3, (1974) 461 – 473. http://www.sciencedirect.com/science/article/pii/0550321374901540.

[17] A. Buckley *et al.*, "General-purpose event generators for LHC physics," *Phys. Rept.* **504** (2011) 145–233, arXiv:1101.2599 [hep-ph].

[18] A. Buckley, J. Ferrando, S. Lloyd, K. Nordström, B. Page, M. Rüfenacht, M. Schönherr, and G. Watt, "LHAPDF6: parton density access in the LHC precision era," *Eur. Phys. J.* **C75** (2015) 132, arXiv:1412.7420 [hep-ph].

[19] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, "MadGraph 5 : Going Beyond," *JHEP* **06** (2011) 128, arXiv:1106.0522 [hep-ph].

[20] J. Baglio *et al.*, "Release Note - VBFNLO 2.7.0," arXiv:1404.3940 [hep-ph].

[21] S. Alioli *et al.*, "Update of the Binoth Les Houches Accord for a standard interface between Monte Carlo tools and one-loop programs," *Comput. Phys. Commun.* **185** (2014) 560–571, arXiv:1308.3462 [hep-ph].

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

[23] T. D. Gottschalk, "BACKWARDS EVOLVED INITIAL STATE PARTON SHOWERS," *Nucl. Phys.* **B277** (1986) 700–738.

[24] S. Plätzer and M. Sjödahl, "Subleading N_c improved Parton Showers," *JHEP* **07** (2012) 042, arXiv:1201.0260 [hep-ph].

[25] Z. Nagy and D. E. Soper, "Effects of subleading color in a parton shower," *JHEP* **07** (2015) 119, arXiv:1501.00778 [hep-ph].

[26] J. Bellm, "Colour Rearrangement for Dipole Showers," *Eur. Phys. J.* **C78** no. 7, (2018) 601, arXiv:1801.06113 [hep-ph].

[27] J. Isaacson and S. Prestel, "On Stochastically Sampling Color Configurations," arXiv:1806.10102 [hep-ph].

[28] S. Gieseke, P. Stephens, and B. Webber, "New formalism for QCD parton showers," *JHEP* **12** (2003) 045, arXiv:hep-ph/0310083 [hep-ph].

[29] S. Plätzer and S. Gieseke, "Coherent Parton Showers with Local Recoils," *JHEP* **01** (2011) 024, arXiv:0909.5593 [hep-ph].

[30] G. Altarelli and G. Parisi, "Asymptotic freedom in parton language," *Nuclear Physics B* **126** no. 2, (1977) 298 – 318. http://www.sciencedirect.com/science/article/pii/0550321377903844.

[31] Y. L. Dokshitzer, "Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics.," *Sov. Phys. JETP* **46** (1977) 641–653. [Zh. Eksp. Teor. Fiz.73,1216(1977)].

[32] V. N. Gribov and L. N. Lipatov, "Deep inelastic e p scattering in perturbation theory," *Sov. J. Nucl. Phys.* **15** (1972) 438–450. [Yad. Fiz.15,781(1972)].

[33] D. Amati and G. Veneziano, "Preconfinement as a property of perturbative qcd," *Physics Letters B* **83** no. 1, (1979) 87 – 92. http://www.sciencedirect.com/science/article/pii/0370269379908967.

[34] G. Marchesini, L. Trentadue, and G. Veneziano, "Space-time description of colour screening via jet calculus techniques," *Nuclear Physics B* **181** no. 2, (1981) 335 – 346. http://www.sciencedirect.com/science/article/pii/0550321381903576.

[35] A. Bassetto, M. Ciafaloni, and G. Marchesini, "Color singlet distributions and mass damping in perturbative qcd," *Physics Letters B* **83** (05, 1979) 207–212.

[36] S. Catani and M. H. Seymour, "A General algorithm for calculating jet cross-sections in NLO QCD," *Nucl. Phys.* **B485** (1997) 291–419, arXiv:hep-ph/9605323 [hep-ph]. [Erratum: Nucl. Phys.B510,503(1998)].

[37] B. Webber, "A qcd model for jet fragmentation including soft gluon interference," *Nuclear Physics B* **238** no. 3, (1984) 492 – 528. http://www.sciencedirect.com/science/article/pii/055032138490333X.

[38] D. Grellscheid and P. Richardson, "Simulation of Tau Decays in the Herwig++ Event Generator," arXiv:0710.1951 [hep-ph].

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

[40] **Particle Data Group** Collaboration, "Review of particle physics," *Phys. Rev. D* **98** (Aug, 2018) 030001. https://link.aps.org/doi/10.1103/PhysRevD.98.030001.

[41] P. Richardson, "Spin correlations in Monte Carlo simulations," *JHEP* **11** (2001) 029, arXiv:hep-ph/0110108 [hep-ph].

[42] I. G. Knowles, "Spin Correlations in Parton - Parton Scattering," *Nucl. Phys.* **B310** (1988) 571–588.

[43] I. G. Knowles, "A Linear Algorithm for Calculating Spin Correlations in Hadronic Collisions," *Comput. Phys. Commun.* **58** (1990) 271–284.

[44] J. C. Collins, "Spin Correlations in Monte Carlo Event Generators," *Nucl. Phys.* **B304** (1988) 794–804.

[45] D. J. Lange, "The evtgen particle decay simulation package," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **462** no. 1, (2001) 152 – 155. http://www.sciencedirect.com/science/article/pii/S0168900201000894. BEAUTY2000, Proceedings of the 7th Int. Conf. on B-Physics at Hadron Machines.

[46] T. Sjöstrand and M. van Zijl, "A Multiple Interaction Model for the Event Structure in Hadron Collisions," *Phys. Rev.* **D36** (1987) 2019.

[47] J. M. Butterworth, J. R. Forshaw, and M. H. Seymour, "Multiparton interactions in photoproduction at HERA," *Z. Phys.* **C72** (1996) 637–646, arXiv:hep-ph/9601371 [hep-ph].

[48] L. Durand and H. Pi, "High-energy nucleon-nucleus scattering and cosmic-ray cross sections," *Phys. Rev. D* **38** (Jul, 1988) 78–84. http://link.aps.org/doi/10.1103/PhysRevD.38.78.

[49] I. Borozan and M. H. Seymour, "An Eikonal model for multiparticle production in hadron hadron interactions," *JHEP* **09** (2002) 015, arXiv:hep-ph/0207283 [hep-ph].

[50] L. Durand and H. Pi, "Semihard qcd and high-energy pp and $\overline{p}p$ scattering," *Phys. Rev. D* **40** (Sep, 1989) 1436–1445. http://link.aps.org/doi/10.1103/PhysRevD.40.1436.

[51] M. Bähr, J. M. Butterworth, and M. H. Seymour, "The Underlying Event and the Total Cross Section from Tevatron to the LHC," *JHEP* **01** (2009) 065, arXiv:0806.2949 [hep-ph].

[52] M. Bähr, J. M. Butterworth, S. Gieseke, and M. H. Seymour, "Soft interactions in Herwig++," in *Proceedings, 1st International Workshop on Multiple Partonic Interactions at the LHC (MPI08): Perugia, Italy, October 27-31, 2008*, pp. 239–248. 2009. arXiv:0905.4671 [hep-ph].

[53] S. Gieseke, F. Loshaj, and P. Kirchgaeßer, "Soft and diffractive scattering with the cluster model in Herwig," *Eur. Phys. J.* **C77** no. 3, (2017) 156, arXiv:1612.04701 [hep-ph].

[54] D. Amati, A. Stanghellini, and S. Fubini, "Theory of high-energy scattering and multiple production," *Il Nuovo Cimento (1955-1965)* **26** no. 5, (1962) 896–954. http://dx.doi.org/10.1007/BF02781901.

[55] M. Baker and K. Ter-Martirosyan, "Gribov's reggeon calculus: Its physical basis and implications," *Physics Reports* **28** no. 1, (1976) 1 – 143. http://www.sciencedirect.com/science/article/pii/0370157376900028.

[56] **ATLAS** Collaboration, G. Aad *et al.*, "Rapidity gap cross sections measured with the ATLAS detector in pp collisions at $\sqrt{s} = 7$ TeV," *Eur. Phys. J.* **C72** (2012) 1926, arXiv:1201.2808 [hep-ex].

[57] S. Gieseke, F. Loshaj, and M. Myska, "Towards Diffraction in Herwig," in *Proceedings, 7th International Workshop on Multiple Partonic Interactions at the LHC (MPI@LHC 2015): Miramare, Trieste, Italy, November 23-27, 2015*, pp. 53–57. 2016. arXiv:1602.04690 [hep-ph].

[58] M. Bähr, *Underlying Event Simulation in the Herwig++ Event Generator*. PhD thesis, 2006. https://www.itp.kit.edu/prep/phd/PSFiles/Diss_Baehr.pdf. Karlsruhe, KIT, Diss., 2008.

[59] A. Donnachie and P. V. Landshoff, "Total cross-sections," *Phys. Lett.* **B296** (1992) 227–232, arXiv:hep-ph/9209205 [hep-ph].

[60] A. Donnachie and P. V. Landshoff, "Does the hard pomeron obey Regge factorization?," *Phys. Lett.* **B595** (2004) 393–399, arXiv:hep-ph/0402081 [hep-ph].

[61] **CDF** Collaboration, F. Abe *et al.*, "Measurement of small angle $\bar{p}p$ elastic scattering at $\sqrt{s} = 546$ GeV and 1800 GeV," *Phys. Rev.* **D50** (1994) 5518–5534.

[62] S. Gieseke, P. Kirchgaeßer, and F. Loshaj, "Soft and diffractive scattering," in *Proceedings, 17th conference on Elastic and Diffractive Scattering (EDS Blois 2017): Prague, Czech Republic, June 26-30, 2017*. 2017. arXiv:1710.10925 [hep-ph].

[63] A. H. Mueller, "O(2,1) Analysis of Single Particle Spectra at High-energy," *Phys. Rev.* **D2** (1970) 2963–2968.

[64] V. Barone and E. Predazzi, *High-Energy Particle Diffraction*, vol. v.565 of *Texts and Monographs in Physics*. Springer-Verlag, Berlin Heidelberg, 2002. http://www-spires.fnal.gov/spires/find/books/www?cl=QC794.6.C6B37::2002.

[65] **ALICE** Collaboration, B. Abelev *et al.*, "Measurement of inelastic, single- and double-diffraction cross sections in proton–proton collisions at the LHC with ALICE," *Eur. Phys. J.* **C73** no. 6, (2013) 2456, arXiv:1208.4968 [hep-ex].

[66] S. Gieseke, "private communication.".

[67] **ATLAS** Collaboration, G. Aad *et al.*, "Charged-particle multiplicities in pp interactions measured with the ATLAS detector at the LHC," *New J. Phys.* **13** (2011) 053033, arXiv:1012.5104 [hep-ex].

[68] Patrick Kirchgaeßer, "Soft Physics at the LHC," 2016. Master's Thesis.

[69] J. Bellm, C. B. Duncan, S. Gieseke, M. Myska, and A. Siódmok, "Spacetime Colour Reconnection in Herwig 7," arXiv:1909.08850 [hep-ph].

[70] S. Ferreres-Solé and T. Sjöstrand, "The space–time structure of hadronization in the Lund model," *Eur. Phys. J.* **C78** no. 11, (2018) 983, arXiv:1808.04619 [hep-ph].

[71] C. Röhr, *Simulation of Soft Inclusive Events at Hadron Colliders*. PhD thesis, 2014. http://digbib.ubka.uni-karlsruhe.de/volltexte/1000039423. Karlsruhe, KIT, Diss., 2014.

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

[73] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equation of State Calculations by Fast Computing Machines," *J. Chem. Phys* **21** no. 6, (1953) 1087–1092.

[74] D. Reichelt, P. Richardson, and A. Siódmok, "Improving the Simulation of Quark and Gluon Jets with Herwig 7," *Eur. Phys. J.* **C77** no. 12, (2017) 876, arXiv:1708.01491 [hep-ph].

[75] **CMS** Collaboration, V. Khachatryan *et al.*, "Strange Particle Production in *pp* Collisions at $\sqrt{s} = 0.9$ and 7 TeV," *JHEP* **05** (2011) 064, arXiv:1102.4282 [hep-ex].

[76] **ALICE** Collaboration, J. Adam *et al.*, "Enhanced production of multi-strange hadrons in high-multiplicity proton-proton collisions," *Nature Phys.* **13** (2017) 535–539, arXiv:1606.07424 [nucl-ex].

[77] J. R. Christiansen and P. Z. Skands, "String Formation Beyond Leading Colour," *JHEP* **08** (2015) 003, arXiv:1505.01681 [hep-ph].

[78] **WA97** Collaboration, E. Andersen *et al.*, "Strangeness enhancement at mid-rapidity in Pb Pb collisions at 158-A-GeV/c," *Phys. Lett.* **B449** (1999) 401–406.

[79] **NA49** Collaboration, H. Appelshäuser *et al.*, "Xi and anti-xi production in 158-GeV / nucleon Pb + Pb collisions," *Phys. Lett.* **B444** (1998) 523–530, arXiv:nucl-ex/9810005 [nucl-ex].

[80] **STAR** Collaboration, B. I. Abelev *et al.*, "Enhanced strange baryon production in Au + Au collisions compared to p + p at s(NN)**(1/2) = 200-GeV," *Phys. Rev.* **C77** (2008) 044908, arXiv:0705.2511 [nucl-ex].

[81] **ALICE** Collaboration, B. B. Abelev *et al.*, "Multi-strange baryon production at mid-rapidity in Pb-Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV," *Phys. Lett.* **B728** (2014) 216–227, arXiv:1307.5543 [nucl-ex]. [Erratum: Phys. Lett.B734,409(2014)].

[82] **ALICE** Collaboration, B. B. Abelev *et al.*, "Multiplicity Dependence of Pion, Kaon, Proton and Lambda Production in p-Pb Collisions at $\sqrt{s_{NN}} = 5.02$ TeV," *Phys. Lett.* **B728** (2014) 25–38, arXiv:1307.6796 [nucl-ex].

[83] **ALICE** Collaboration, J. Adam *et al.*, "Multi-strange baryon production in p-Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV," *Phys. Lett.* **B758** (2016) 389–401, arXiv:1512.07227 [nucl-ex].

[84] K. Werner, B. Guiot, I. Karpenko, and T. Pierog, "Analysing radial flow features in p-Pb and p-p collisions at several TeV by studying identified particle production in EPOS3," *Phys. Rev.* **C89** no. 6, (2014) 064903, arXiv:1312.1233 [nucl-th].

[85] A. Ortiz Velasquez, P. Christiansen, E. Cuautle Flores, I. A. Maldonado Cervantes, and G. Paić, "Color reconnection and flowlike patterns in *pp* collisions," *Phys. Rev. Lett.* **111** (Jul, 2013) 042001. https://link.aps.org/doi/10.1103/PhysRevLett.111.042001.

[86] C. Bierlich, G. Gustafson, L. Lönnblad, and A. Tarasov, "Effects of Overlapping Strings in pp Collisions," *JHEP* **03** (2015) 148, arXiv:1412.6259 [hep-ph].

[87] C. Bierlich and J. R. Christiansen, "Effects of color reconnection on hadron flavor observables," *Phys. Rev.* **D92** no. 9, (2015) 094010, arXiv:1507.02091 [hep-ph].

[88] N. Fischer and T. Sjöstrand, "Thermodynamical String Fragmentation," *JHEP* **01** (2017) 140, arXiv:1610.09818 [hep-ph].

[89] **ALICE** Collaboration, K. Aamodt *et al.*, "Charged-particle multiplicity measurement in proton-proton collisions at $\sqrt{s} = 7$ TeV with ALICE at LHC," *Eur. Phys. J.* **C68** (2010) 345–354, arXiv:1004.3514 [hep-ex].

[90] **ALICE** Collaboration, J. Adam *et al.*, "Measurement of pion, kaon and proton production in proton–proton collisions at $\sqrt{s} = 7$ TeV," *Eur. Phys. J.* **C75** no. 5, (2015) 226, arXiv:1504.00024 [nucl-ex].

[91] A. Buckley, J. Butterworth, L. Lönnblad, D. Grellscheid, H. Hoeth, J. Monk, H. Schulz, and F. Siegert, "Rivet user manual," *Comput. Phys. Commun.* **184** (2013) 2803–2819, arXiv:1003.0694 [hep-ph].

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

[93] A. Buckley, H. Hoeth, H. Schulz, and J. E. von Seggern, "Monte Carlo event generator validation and tuning for the LHC," *PoS* **ACAT08** (2008) 112, arXiv:0902.4403 [hep-ph].

[94] J. Bellm *et al.*, "Herwig 7.2 Release Note," arXiv:1912.06509 [hep-ph].

[95] T. Sjöstrand and V. A. Khoze, "On Color rearrangement in hadronic W+ W- events," *Z. Phys.* **C62** (1994) 281–310, arXiv:hep-ph/9310242 [hep-ph].

[96] L. Lönnblad, "Reconnecting colored dipoles," *Z. Phys.* **C70** (1996) 107–114.

[97] J. R. Christiansen and T. Sjöstrand, "Color reconnection at future e^+ e^- colliders," *Eur. Phys. J.* **C75** no. 9, (2015) 441, arXiv:1506.09085 [hep-ph].

[98] S. Plätzer, "Summing Large-N Towers in Colour Flow Evolution," *Eur. Phys. J.* **C74** no. 6, (2014) 2907, arXiv:1312.2448 [hep-ph].

[99] R. A. Martinez, M. De Angelis, J. R. Forshaw, S. Plätzer, and M. H. Seymour, "Soft gluon evolution and non-global logarithms," arXiv:1802.08531 [hep-ph].

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

[101] T. Becher and M. Neubert, "Infrared singularities of scattering amplitudes in perturbative QCD," *Phys. Rev. Lett.* **102** (2009) 162001, arXiv:0901.0722 [hep-ph]. [Erratum: Phys. Rev. Lett.111,no.19,199905(2013)].

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

[103] M. Arioli, B. Codenotti, and C. Fassino, "The padé method for computing the matrix exponential," *Linear Algebra and its Applications* **240** (1996) 111 – 130. http://www.sciencedirect.com/science/article/pii/0024379594001901.

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

[105] S. Jadach, "Rapidity generator for monte-carlo calculations of cylindrical phase space," *Computer Physics Communications* **9** no. 5, (1975) 297 – 304. http://www.sciencedirect.com/science/article/pii/0010465575900818.

[106] G. Alner et al., "The UA5 high energy pp simulation program," *Nuclear Physics B* **291** (1987) 445 – 502. http://www.sciencedirect.com/science/article/pii/0550321387904810.

[107] C. Bierlich, G. Gustafson, and L. Lönnblad, "A shoving model for collectivity in hadronic collisions," arXiv:1612.05132 [hep-ph].

[108] T. Pierog, I. Karpenko, J. M. Katzy, E. Yatsenko, and K. Werner, "EPOS LHC: Test of collective hadronization with data measured at the CERN Large Hadron Collider," *Phys. Rev.* **C92** no. 3, (2015) 034906, arXiv:1306.0121 [hep-ph].

[109] B. Blok, C. D. Jäkel, M. Strikman, and U. A. Wiedemann, "Collectivity from interference," *JHEP* **12** (2017) 074, arXiv:1708.08241 [hep-ph].

[110] P. Skands, S. Carrazza, and J. Rojo, "Tuning PYTHIA 8.1: the Monash 2013 Tune," *Eur. Phys. J.* **C74** no. 8, (2014) 3024, arXiv:1404.5630 [hep-ph].

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

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

[113] **SLD** Collaboration, K. Abe *et al.*, "Production of pi+, pi-, K+, K-, p and anti-p in light (uds), c and b jets from Z0 decays," *Phys. Rev.* **D69** (2004) 072003, arXiv:hep-ex/0310017 [hep-ex].

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

[115] M. Azarkin, P. Kotko, A. Siódmok, and M. Strikman, "Studying minijets and MPI with rapidity correlations," *Eur. Phys. J.* **C79** no. 3, (2019) 180, arXiv:1806.09016 [hep-ph].

[116] **UA5** Collaboration, G. J. Alner *et al.*, "The UA5 High-Energy anti-p p Simulation Program," *Nucl. Phys.* **B291** (1987) 445–502.

[117] J. Bellm, S. Gieseke, and P. Kirchgaeßer, "Improving the description of multiple interactions in Herwig," arXiv:1911.13149 [hep-ph].

[118] **CMS** Collaboration, V. Khachatryan *et al.*, "Measurement of diffraction dissociation cross sections in pp collisions at $\sqrt{s} = 7$ TeV," *Phys. Rev.* **D92** no. 1, (2015) 012003, arXiv:1503.08689 [hep-ex].

[119] **STAR** Collaboration, B. I. Abelev *et al.*, "Systematic Measurements of Identified Particle Spectra in pp, d^+ Au and Au+Au Collisions from STAR," *Phys. Rev.* **C79** (2009) 034909, arXiv:0808.2041 [nucl-ex].

[120] **UA1** Collaboration, C. Albajar *et al.*, "A Study of the General Characteristics of $p\bar{p}$ Collisions at $\sqrt{s} = 0.2$-TeV to 0.9-TeV," *Nucl. Phys.* **B335** (1990) 261–287.

[121] **ATLAS** Collaboration, G. Aad *et al.*, "Measurement of underlying event characteristics using charged particles in pp collisions at $\sqrt{s} = 900GeV$ and 7 TeV with the ATLAS detector," *Phys. Rev.* **D83** (2011) 112001, arXiv:1012.0791 [hep-ex].

[122] **ATLAS** Collaboration, G. Aad *et al.*, "Measurements of underlying-event properties using neutral and charged particles in pp collisions at 900 GeV and 7 TeV with the ATLAS detector at the LHC," *Eur. Phys. J.* **C71** (2011) 1636, arXiv:1103.1816 [hep-ex].

[123] **CDF** Collaboration, F. Abe *et al.*, "Transverse Momentum Distributions of Charged Particles Produced in $\bar{p}p$ Interactions at $\sqrt{s} = 630$ GeV and 1800 GeV," *Phys. Rev. Lett.* **61** (1988) 1819.

[124] **CDF** Collaboration, F. Abe *et al.*, "Pseudorapidity distributions of charged particles produced in $\bar{p}p$ interactions at $\sqrt{s} = 630$ GeV and 1800 GeV," *Phys. Rev.* **D41** (1990) 2330. [119(1989)].

[125] **CDF** Collaboration, D. Acosta *et al.*, "Soft and Hard Interactions in $p\bar{p}$ Collisions at $\sqrt{s} = 1800$-GeV and 630-GeV," *Phys. Rev.* **D65** (2002) 072005.

[126] **CDF** Collaboration, T. Affolder *et al.*, "Charged jet evolution and the underlying event in proton-antiproton collisions at 1.8 TeV," *Phys. Rev.* **D65** (2002) 092002.

[127] **CDF** Collaboration, D. E. Acosta *et al.*, "Underlying event in hard interactions at the Fermilab Tevatron $\bar{p}p$ collider," *Phys. Rev.* **D70** (2004) 072002, hep-ex/0404004.

[128] **ATLAS** Collaboration, M. Aaboud *et al.*, "Measurement of charged-particle distributions sensitive to the underlying event in $\sqrt{s} = 13$ TeV proton-proton collisions with the ATLAS detector at the LHC," *JHEP* **03** (2017) 157, arXiv:1701.05390 [hep-ex].

[129] **ATLAS** Collaboration, M. Aaboud *et al.*, "Charged-particle distributions at low transverse momentum in $\sqrt{s} = 13$ TeV pp interactions measured with the ATLAS detector at the LHC," *Eur. Phys. J.* **C76** no. 9, (2016) 502, arXiv:1606.01133 [hep-ex].

[130] **CMS** Collaboration, A. M. Sirunyan *et al.*, "Measurement of charged particle spectra in minimum-bias events from proton–proton collisions at $\sqrt{s} = 13\,\text{TeV}$," *Eur. Phys. J.* **C78** no. 9, (2018) 697, arXiv:1806.11245 [hep-ex].

[131] **CMS** Collaboration, S. Chatrchyan *et al.*, "Measurement of differential cross sections for the production of a pair of isolated photons in pp collisions at $\sqrt{s} = 7\,\text{TeV}$," *Eur. Phys. J.* **C74** no. 11, (2014) 3129, arXiv:1405.7225 [hep-ex].

[132] C. O. Rasmussen and T. Sjöstrand, "Models for total, elastic and diffractive cross sections," *Eur. Phys. J.* **C78** no. 6, (2018) 461, arXiv:1804.10373 [hep-ph].

[133] J. Bellm and C. Bierlich, "PISTA: Posterior Ion STAcking," arXiv:1807.01291 [hep-ph].