LU TP 18-07
May 2018

THE SPACE–TIME STRUCTURE OF HADRON PRODUCTION IN HIGH-ENERGY COLLISIONS

Silvia Ferreres-Solé

Department of Astronomy and Theoretical Physics, Lund University

Master thesis supervised by Torbjörn Sjöstrand

LUNDS UNIVERSITET

Abstract

In this thesis, the space–time picture of the hadronization model has been implemented in the event generator Pythia. From this implementation, the hadronic behaviour has been studied for different scenarios, in order to begin addressing the unexpected collective behaviour of high-multiplicity pp events, which is typical for heavy-ion collisions. First, the temporal and radial evolution of pp and toy $q\overline{q}$ systems are analysed, illustrating the general motion of hadrons in the two systems. Then, the hadronic density is studied in the central region of the collisions, showing the presence of hadronic close-packing. This issue is further approached by tackling the hadronic density for different multiplicities, where pp collisions with higher multiplicity exhibit more significant close-packing than low-multiplicity collisions. It turns out, however, that colour reconnection gives a larger transverse area over which particle production occurs, and thereby a lesser density than without it. Finally, the hadronic overlap is studied as a function of transverse momentum for each newly generated hadron and for different multiplicities, illustrating that, although the importance of close-packing increases with multiplicity, it mainly affects low-p_T hadrons.

Populärvetenskaplig sammanfattning

Partikelkolliderare är stora maskiner som accelererar partiklar till höga energier, för att sedan låta dem kollidera. Därvid bildas ett stort antal produkter, som kan spåras i detektorer. De processer som inträffar under och efter kollisionerna beskrivs huvudsakligen av standardmodellen, en teori som klassificerar de fundamentala partiklarna och växelverkningarna mellan dessa. Även om standardmodellen beskriver observerade fenomen rättså precist, finns det experimentella observationer som inte kan förklaras genom denna modell, såsom gravitation. Därför utvecklas nya teorier som kompletterar standardmodellen, och som är kapabla att inkludera sådana oförklarade observationer.

Den forskning som görs vid partikelkolliderare består av att testa olika teorier genom att jämföra experimentella data med förutsägelser från dessa teorier, erhållna genom simuleringar. Sammansättningen av sådana simuleringar i ett datorprogram kallas för en händelsegenerator. Sålunda simulerar en händelsegenerator partikelkollisioner och de produkter som genereras av dem, vilka kan vara slutliga stabila partiklar, eller sönderfalla till nya partiklar, eller växelverka med andra partiklar. I just detta projekt fokuserar vi på Pythia, en händelsegenerator som utvecklats i Lund.

Händelsegeneratorer är den hörnsten som relaterar teoretisk och experimentell partikelfysik. De måste samtidigt reproducera det observerade experimentella uppförandet och förutsägelserna från den teoretiska modell som skall testas. Sålunda är utvecklingen av dem en komplicerad och teknisk uppgift, som måste ta hänsyn till olika modeller och ett stort antal mindre detaljer. Att bekräfta eller förkasta en teori beror till stor del av händelsegeneratorernas resultat. Därför måste de simuleringar som genomföres vara så noggranna och kompletta som möjligt. I denna avhandling introduceras en ny komponent till Pythia, som ger information om de positioner vid vilka partiklar skapas efter kollisionerna. Denna nya kod kommer att bredda den information som finns om partiklarna som skapas i acceleratorkollisioner, och sålunda hjälpa till att förstå de processer som styr uppförandet hos materiens minsta beståndsdelar.

Contents

1 Introduction **3**

2 LHC physics overview **4**

3 The Lund String Model **6**
3.1 The linear force field in QCD 6
3.2 Two-parton system 7
3.2.1 String motion 7
3.2.2 String breaking and hadron formation 9
3.2.3 How are the breakup points obtained? 11
3.2.4 Tunneling process and transverse momentum p_T 14
3.2.5 Massive quarks 14
3.3 Multiparton systems 16
3.4 Lund String Model implementation 21

4 Space–time implementation **22**
4.1 Simplest $q\overline{q}$ system and hadron production points 22
4.2 More complex topologies 23
4.3 Gluon–loops 25
4.4 Smearing in transverse space 26
4.5 Massive quark implementation 27
4.6 Other implementation details 30

5 Hadron density studies **31**
5.1 Longitudinal and transverse distributions 31
5.2 Temporal and radial evolution of hadron production 34
5.3 Close-packing of hadron production in the central region 38
5.4 Hadron production at different multiplicities 40
5.5 Close-packing analysis in hadron rest frame 45

6 Summary **48**

A Space–time location of the final breakup **49**

B Correction to non-physical situations **54**

1 Introduction

Important progress has been achieved in the understanding of the Universe and matter formation over the past decades. Part of this progress has been possible thanks to particle accelerators, such as the Large Hadron Collider (LHC), which accelerate particles to high energies to make them collide afterwards. During the collisions, some of the conditions of the early Universe are reproduced and studied by analysing the particles generated. These particles are tracked by detectors, collecting the experimental data that is later analysed and used to test different theories, such as the Standard Model.

The Standard Model is the theoretical scheme that describes and classifies the elementary particles and the interactions among them [1, 2]. Although this model describes accurately the experimental phenomena, it cannot explain some experimental observations, such as Dark Matter [1]. Hence, new theories complementing the Standard Model are needed to address those unexplained aspects. These theories, known as Beyond the Standard Model (BSM) theories [1], can also be tested from the analysis of the data collected in particle colliders.

In order to test different theories, experimental data is compared with theoretical predictions, reproduced through computer simulations using Monte Carlo techniques. The ensemble of such simulations is collected into an event generator [3]. Since event generators are the cornerstone relating experiment and theory, their simulations must reproduce the experimental observations while obeying the principles of the theories to be tested [3, 4]. The three main event generators currently used are Pythia [5, 6], Herwig [7, 8] and Sherpa [9]. This thesis focuses on the first, in which particle collisions are simulated at the partonic level, composed of quarks and gluons, whose interactions are described by the Quantum Chromodynamics (QCD) theory [2]. These partons combine to create hadrons in a process known as hadronization, which is the main focus of this thesis, modelled in Pythia by the Lund String Model [10, 11].

The LHC collides mainly protons and heavy-ions, such as lead nuclei. One of the main differences between the two is the creation, in heavy-ion collisions, of Quark-Gluon Plasma (QGP), a hot and dense state of matter composed of free partons [12]. The formation of this plasma requires a system formed by big nuclei that lives long enough to reach thermal equilibrium. The *pp* systems are not so long-lived as to generate QGP, a fact that has been supported by the different behaviour of hadron production in heavy-ion and *pp* collisions [12]. Nevertheless, this does not seem to be the case for high-multiplicity *pp* collisions, since an unexpected behaviour of the hadron production has been observed in those collisions over the last few years. Particularly, high-multiplicity *pp* collisions show a collective behaviour typical of heavy-ion collisions, such as the enhancement of strangeness production detected in ALICE [13] or the long-range azimuthal correlations in CMS [14, 15] and ATLAS [16]. These observations lead to two open scenarios: either the plasma is also generated in smaller systems than predicted, or else the signals used to define the presence of QGP and its formation can be questioned.

The aim of this thesis is to begin addressing this unexpected behaviour by studying the space–time hadronic density in different pp collisions. A general overview of LHC physics and event generators can be found in section 2. So far, the hadronization process in Pythia was only developed in an energy–momentum picture, modelled by the Lund String Model, presented in section 3. Hence, the space–time picture is firstly implemented in Pythia, as described in section 4. From that implementation, the space–time location at which hadrons are created during hadronization is determined, and the space–time hadronic density is analysed in different situations, as presented in section 5.

Note that, in this thesis, natural units are assumed, i.e. $c = \hbar = 1$, simplifying relations among variables.

2 LHC physics overview

Quantum Chromodynamics (QCD) is the theory that characterizes the strong interactions among coloured particles, i.e. quarks and gluons [17]. Those interactions are mediated by gluons, the force carries of QCD. Contrary to photons in Quantum Electrodynamics (QED), gluons carry colour charge, which implies that they are self-interacting [1]. QCD constitutes, along with QED and the weak interactions, the cornerstones of the Standard Model of particle physics [2].

One of the two fundamental parameters of QCD is the strong coupling constant, α_s, which determines the strength of the interaction in a specific process [2]. Due to the running of α_s, two different limits are defined in the theory, which have to be addressed using opposite approaches [17]. On the one hand, the value of α_s is small for processes with high momentum transfer. Those processes are tackled using perturbation theory, since they can be formulated by adding higher-order terms to give an expansion that, hopefully, converges rapidly. On the other hand, processes with low momentum transfer are related to a strong coupling, where non-perturbative approaches have to be followed [17]. In such cases, the coupling among partons is so strong that confinement arises, forming hadrons.

Although all the QCD processes take place in each particle collision at the LHC, only the confined states are tracked by the detectors, as partons are already confined in hadrons when hitting the different detector components. By interpreting and understanding the experimental data and the performance of the detector, conclusions can be drawn about QCD and its properties. Although the hard QCD processes can be determined using perturbation theory, some of the inputs needed in the calculation, such as the exact parton distribution functions of the initial protons, are undefined in the theory. Several techniques have been developed to obtain such distributions by performing quantitative predictions on the perturbative QCD level [17]. Moreover, the understanding of the whole QCD picture needs the development of some models that attempt to describe the mechanisms in non-perturbative QCD. All those calculations and processes are carried out and modelled in an event generator, making the comparison of theory with data possible.

An event taking place at the LHC is described and simulated using several mechanisms. To illustrate this procedure, we will consider pp collisions, since this thesis is focused on them. First, the two proton bunches are accelerated to the highest possible energy, when they are made to collide. Due to Lorentz contraction, protons acquire a pancake shape when accelerated to high energies in the LHC. Because protons are composite particles formed by partons, when they collide, partons from each incoming beam interact, generating several binary partonic collisions. Each of those interactions, known as multiparton interaction (MPI) [18], takes place separately, without interfering with other interactions. For a particular study, one may want to single out those events that contain a specific process, such as $u\overline{d} \longrightarrow W^+$. Often these processes have a higher momentum transfer than the average interactions and are, therefore, called hard interactions [4]. Those processes can be described by the Standard Model or by BSM theories.

During the hard interaction, some resonances such as W or Z^0 bosons, with short lifetimes, can be produced. Their decays are regarded as part of the hard interaction, since their decay products carry important information for the process such as spin correlations [18]. Apart from the hard process itself, partons can also be emitted due to bremsstrahlung, generated by both accelerated colour and electromagnetic charges [18]. If the radiated partons are associated with the incoming colliding partons, the process is known as Initial-State Radiation (ISR) while, if related to the outgoing partons after the collision, the emission is known as Final-Sate Radiation (FSR) [4]. Both ISR and FSR are simulated in Pythia using the parton shower approach.

Up until this point, confinement forces have not been significant, because all the processes described so far are part of the QCD short-distance perturbative regime. This changes when the partons created at previous stages separate from each other, decreasing the momentum transfer and, thus, increasing the value of α_s. Due to the increasing importance of confinement, further partons are created non-perturbatively and together combine to create hadrons. This process, called hadronization [3], is modelled in Pythia using the Lund String Model, presented in the next section.

Once the hadrons have been created, they might decay through strong, electromagnetic or weak interactions. The lifetimes of the particles that decay are related to the values of the coupling constant of the interaction ruling the decay. Since the strong coupling constant of processes with low momentum transfer is large, strong decays are the first to take place. An example of such decays is the ρ^0 decay $\rho^0 \to \pi^+\pi^-$, involving the exchange of at least one gluon [1]. This decay is also possible by the exchange of a photon, in which case it would be an electromagnetic decay [1]. However, when kinematically possible, the strong decay modes always dominate over electromagnetic, because $\alpha_{\text{EM}} < \alpha_s$ [1]. Although $\alpha_{\text{weak}} > \alpha_{\text{EM}}$, electromagnetic decays prevail over weak decays since the latter arise from the exchange of a $W^\pm$ or Z^0 boson, whose masses are quite large, while the photon, the force carrier of electromagnetic processes, is massless [2]. While particles decaying via weak decays are long-lived particles which can travel several thousand meters before decaying,

particles decaying strongly travel only a few fermi before decaying, making the time range of the decays quite wide [22].

The hadrons created in the hadronization process are referred to as primary hadrons and might decay via one of the three decay modes presented, while the hadrons generated from particle decays are referred to as secondary hadrons. The production points of the decay products resulting from strong decays are considered by experimentalists as primary vertices, since the distance travelled by the mother hadron before decaying is of order of only a few fermi [22]. Although only the proton is a stable hadron, all the hadrons that do not decay inside the detector are regarded as final hadrons in an event generator. This selection depends on the lifetime of each hadron, which is related to their decay modes. Hence, only some particles decaying weakly are regarded as stable, namely $\mu^{\pm}$, $\pi^{\pm}$, $K^{\pm}$, K_L and n. Those particles, along with the truly stable particles ($e^{\pm}$, p, $\overline{p}$, γ and ν) [22], are considered final and stable particles in event generator language.

3 The Lund String Model

3.1 The linear force field in QCD

One of the fundamental properties of QCD is the non-appearance of isolated partons. This phenomenon, known as colour confinement, is generated by a potential between a quark and an antiquark approximately of the form [1, 2]

$$V_{\mathrm{QCD}}(r) = -\frac{4}{3}\frac{\alpha_s}{r} + \kappa r \tag{3.1}$$

where r is the distance between q and $\overline{q}$ and α_s, the strong coupling constant.

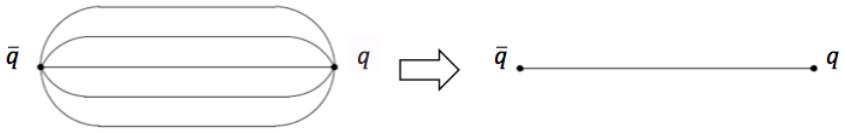


Figure 1: Tubelike potential and string approximation.

The first term of eq. (3.1) is similar to the QED Coulomb potential, apart from the factor of 4/3, which accounts for the normalization of colour charges. The second term corresponds to a linear force field, where κ is a constant. For large distances, the linear term predominates, since the behaviour of the potential field between q and $\overline{q}$ for those distances is tubelike, meaning that the colour field between them is compressed into a tube. As a mathematical approximation, the potential can be expressed as strings or tubes of zero width, as shown in Figure 1. The strings, whose string tension is $\kappa \approx 1$ GeV/fm [10], are responsible for confinement. The existence of this linear confining potential and the value of the string tension have been demonstrated by lattice QCD calculations.

3.2 Two-parton system

3.2.1 String motion

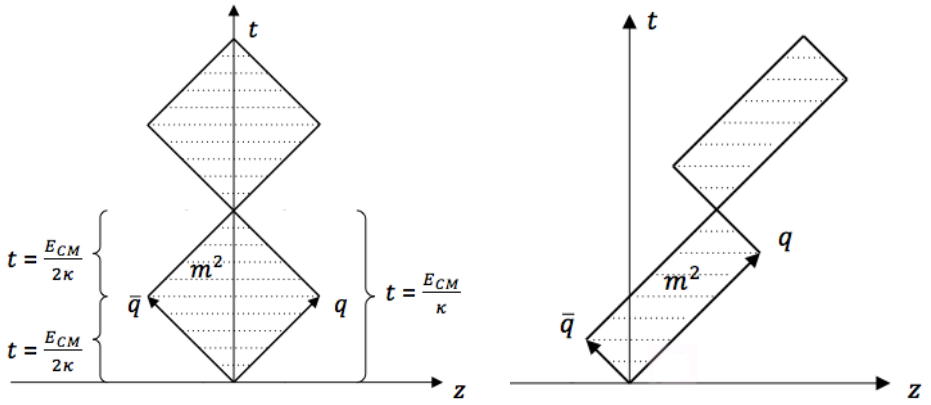
The mathematical approximation of the linear potential as strings is the cornerstone of the Lund String model, used to simulate the hadronization process. To illustrate how the model works, consider the collinear motion in 1+1 dimensional space–time, in the (z, t) coordinates, of a massless quark–antiquark pair with four-momenta p_q and $p_{\overline{q}}$, respectively. For large distances, only accounting for the linear potential, the Hamiltonian of the $q\overline{q}$ system is [10]

$$H = E_q + E_{\overline{q}} + \kappa|z_q - z_{\overline{q}}| \tag{3.2}$$

where $|z_q - z_{\overline{q}}|$ is the distance between q and $\overline{q}$, while E_q and $E_{\overline{q}}$ are the energies of the massless q and $\overline{q}$, implying that $E_{q/\overline{q}} = |\vec{p}_{q/\overline{q}}| = |p_{zq/\overline{q}}|$. From the Hamiltonian, the equation of motion gives rise to a linear relation between the space–time and the energy–momentum pictures

$$\frac{dp_{zq/\overline{q}}}{dt} = \frac{dp_{zq/\overline{q}}}{dz} = \frac{dE_{q/\overline{q}}}{dt} = \frac{dE_{q/\overline{q}}}{dz} = \pm\kappa \tag{3.3}$$

where $p_{zq/\overline{q}}$ stands for the momentum in the z direction of the quark or the antiquark, while the sign in κ accounts for the action of the string and the direction of motion of the parton. When the parton moves along the $+z$ direction, the sign of κ is negative if the string takes energy from the parton and positive otherwise. On the other hand, if the parton moves along the $-z$ direction, the sign of κ is negative when the string gives energy to the parton and positive otherwise.



(a) $q\overline{q}$ system in the CM frame.

(b) $q\overline{q}$ system in a boosted frame with respect to the CM frame.

Figure 2: The motion of a $q\overline{q}$ system, with massless q and $\overline{q}$. Figures adapted from [10].

The linear relation between space–time and energy–momentum is essential to explain the time evolution of a partonic system. To understand why, we will consider the same $q\overline{q}$ system as before, with massless q and $\overline{q}$, whose dynamical behaviour can be illustrated using the yo-yo model [10]. In this model, the q and $\overline{q}$ oscillate back and forth in one-dimensional space due to the influence of the string tension, as depicted in Figure 2a for

the system in the centre of mass (CM) frame. At time $t = 0$, the four-momenta of the q, $\overline{q}$ and the string in the CM frame can be expressed as

$$\begin{aligned} p_{q/\overline{q}}\,(0) &= \frac{E_{\mathrm{CM}}}{2}\,(1;\ 0,\ 0,\ \ \pm 1) \\ p_{\mathrm{string}}\,(0) &= (0;0,0,0) \end{aligned} \tag{3.4}$$

where E_{CM} is the CM energy of the system, the plus sign stands for q and the minus sign for $\overline{q}$, since q moves in the $+z$ direction and $\overline{q}$ along the $-z$ axis. As time passes, q and $\overline{q}$ lose energy, which is transferred to the string due to energy–momentum conservation. Therefore, the four-momenta of the q, $\overline{q}$ and the string evolve with time as

$$\begin{aligned} p_{q/\overline{q}}\,(t) &= \left(\frac{E_{\mathrm{CM}}}{2} - \kappa t\right)(1;\ 0,\ 0,\ \ \pm 1) \\ p_{\mathrm{string}}\,(t) &= (2\kappa t;\ 0,\ 0,\ 0)\,. \end{aligned} \tag{3.5}$$

At time $t = E_{\mathrm{CM}}/2\kappa$, when all the energy of the system is carried by the string, a quarter of the period has passed. At this point, the string starts to lose energy, which is absorbed by the q and the $\overline{q}$. This process is the inverse of the one followed in the first quarter of the oscillation, with the q moving along the $-z$ axis while the $\overline{q}$ moves in the $+z$ direction. Half of the period is reached at time $t = E_{\mathrm{CM}}/\kappa$, when the quarks meet again with their momenta switched

$$\begin{aligned} p_{q/\overline{q}}\left(\frac{E_{\mathrm{CM}}}{\kappa}\right) &= \frac{E_{\mathrm{CM}}}{2}\,(1;\ 0,\ 0,\ \ \mp 1) \\ p_{\mathrm{string}}\left(\frac{E_{\mathrm{CM}}}{\kappa}\right) &= (0;0,0,0)\,. \end{aligned} \tag{3.6}$$

The second half of the full period is a repetition of the first half, only with the role of q and $\overline{q}$ interchanged.

In order to study the yo-yo system further, the light-cone coordinates are the most convenient choice, both in energy–momentum, $\tilde{p}^{\pm} = E \pm p_z$, and in space–time, $\tilde{x}^{\pm} = t \pm z$. Using those coordinates, the invariant mass of the $q\overline{q}$ system can be written as

$$m^2 = (E + p_z)\,(E - p_z) = \tilde{p}^{+}\tilde{p}^{-}. \tag{3.7}$$

If the invariant mass of the system is consistent with a hadron mass, the system will be used to describe that hadron, as long as its flavour content matches the flavour content of the $q\overline{q}$ system [10, 11]. On the contrary, if the invariant mass of the system is greater than the masses of all hadrons with the correct flavour content, the system fragments into different subsystems. This process, known as fragmentation process, is presented in section 3.2.2.

The simplest yo-yo system can be generalized as illustrated in Figure 2b, where the quark and the antiquark have different initial energies, $E_q \neq E_{\overline{q}}$. Equivalently, this system

can be viewed as a boosted copy of the rest–frame setup in Figure 2a. The energy–momentum and space–time coordinates suffer simultaneous transformations under a longitudinal boost and eq. (3.3) holds also after the boost. It is important to point out a seeming contradiction related to longitudinal boosts. Although the string should acquire momentum along the z direction if the system is longitudinally boosted, it only carries energy when the boosted system is studied as it has been done earlier in this section. This apparent contradiction is a consequence of the string being an extended object, with no net longitudinal momentum when both ends are simultaneous. If that string is longitudinally boosted, both string ends will no longer be simultaneous, acquiring the string a net longitudinal momentum. It is only when both ends are simultaneous in the new frame that the string will have no longitudinal momentum, which is the assumption made when addressing similar systems to the one shown in Figure 2b. Hence, the potential and the characteristics of the string, such as the string tension, are Lorentz invariant under longitudinal boosts.

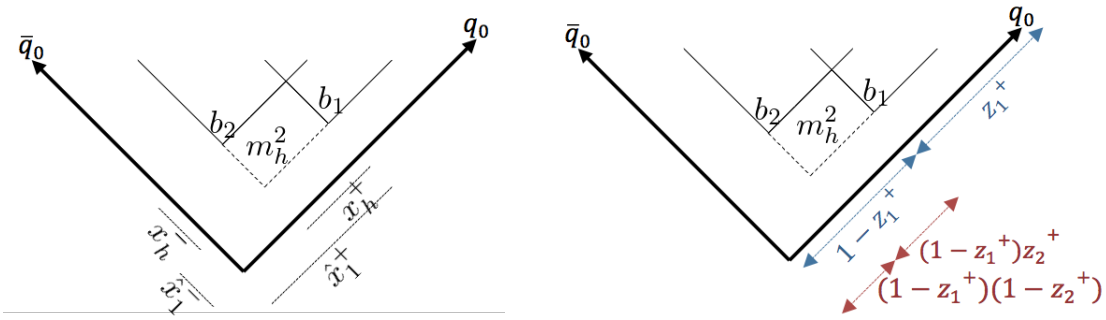
3.2.2 String breaking and hadron formation

In the previous section, the potential between a quark and an antiquark was introduced, illustrating that the potential energy increases with the separation between q_0 and $\overline{q}_0$, making the creation of a new $q_1\overline{q}_1$ pair energetically favourable [1, 2]. In the Lund String Model, pair production breaks the string stretched between q_0 and $\overline{q}_0$ into two string pieces, by the generation of a new $q_1\overline{q}_1$ pair. Naively, the probability for the string to break increases with time, because the string is longer at later times. On the other hand, a break can inhibit later breaks, since it fragments the string into two smaller systems, leaving an in-between region without string, and thus without the ability to break any further. This "exponential decay" behaviour of the string is similar to the one observed in radioactive decay. When the quark and the antiquark pair is generated, both partons have to be created in the same space–time location due to flavour conservation, if transverse momentum is ignored and the quarks are massless.

To illustrate string breaking, consider a $q_0\overline{q}_0$ system in which q_0 and $\overline{q}_0$ are anti-collinear. At some point in space–time, the system breaks into two subsystems by the creation of a new $q_1\overline{q}_1$ pair, such that the anticolour of $\overline{q}_1$ corresponds to the colour of q_0 and similarly for q_1 and $\overline{q}_0$, forming two separate colour singlets [10]. After the break, there are two string segments corresponding to two subsystems: the $q_0\overline{q}_1$ subsystem and the $q_1\overline{q}_0$ subsystem, each of which can either form a hadron or fragment further [10].

A simple $q\overline{q}$ system, formed by massless q and $\overline{q}$, with two breaks, b_1 and b_2, is depicted in Figures 3a and 3b, ignoring the transverse momentum. In such system, the $q_1\overline{q}_2$ subsystem forms a hadron of mass m_h. Due to the linearity between energy–momentum and space–time (eq. (3.3)), the energy and the momentum of the hadrons can be expressed in terms of the space–time coordinates. If considering (t_1, z_1) and (t_2, z_2) to be the space–time coordinates of breaks b_1 and b_2, respectively, the energy and momentum of the hadron formed by $q_1\overline{q}_2$ are [10]

$$E_h = \kappa\,(z_1 - z_2)\,,\; p_{zh} = \kappa\,(t_1 - t_2)\,. \tag{3.8}$$



(a) $\hat{x}^{\pm}$ and $x^{\pm}$ fractions (b) $z^{\pm}$ fractions

Figure 3: Simple $q\overline{q}$ system, where q and $\overline{q}$ are massless, with two breaks, b_1 and b_2. The light-cone coordinates are normalized to unity.

To study the hadron production in the energy-momentum picture, the fractions $\hat{x}^{\pm}$ and $x^{\pm}$ illustrated in Figure 3a are extremely useful. The former define the breakup location in the energy–momentum picture, since they correspond to the absolute coordinates of the breakups when the light-cone coordinates $\tilde{p}^{\pm}$ are normalized to unity. Conversely, the $x^{\pm}$ fractions are momentum fractions related to the hadron, determined as $x_h^{\pm} = \tilde{p}_h^{\pm}/\tilde{p}^{\pm}$, where $\tilde{p}^{+} = E_{q_0} + p_{zq_0}$ and $\tilde{p}^{-} = E_{\overline{q}_0} - p_{z\overline{q}_0}$ are associated to the initial quark and antiquark, respectively. Hence, the $x^{\pm}$ fractions can be understood as the separation between the two adjacent breakup locations, given by $\hat{x}^{\pm}$. From the definitions, the $x^{\pm}$ and $\hat{x}^{\pm}$ fractions are such that $0 \leq x^{\pm} \leq 1$ and $0 \leq \hat{x}^{\pm} \leq 1$. Since hadrons are formed from the partons created in two adjacent breakup points, the energy and momentum of the hadron can be expressed in term of $x^{\pm}$ and $\hat{x}^{\pm}$ as [11]

$$\begin{aligned} E_h &= |\hat{x}_1^+ - \hat{x}_2^+|E_{q_0} + |\hat{x}_1^- - \hat{x}_2^-|E_{\overline{q}_0} = x_h^+ E_{q_0} + x_h^- E_{\overline{q}_0} \\ p_{zh} &= |\hat{x}_1^+ - \hat{x}_2^+|p_{zq_0} + |\hat{x}_1^- - \hat{x}_2^-|p_{z\overline{q}_0} = x_h^+ p_{zq_0} + x_h^- p_{z\overline{q}_0} \end{aligned} \tag{3.9}$$

where the sub-indices in the $\hat{x}^{\pm}$ fractions indicate the string breakups taken into account when forming the hadron. Note that $E_q = p_q$ and $E_{\overline{q}_0} = -p_{\overline{q}_0}$ as the quarks are massless. Using four-vector notation in eq. (3.9), the hadron four-momentum is given by

$$p_h = |\hat{x}_1^+ - \hat{x}_2^+|p_{q_0} + |\hat{x}_1^- - \hat{x}_2^-|p_{\overline{q}_0} = x_h^+ p^+ + x_h^- p^- \tag{3.10}$$

where $p_{q_0} = p^+$ and $p_{\overline{q}_0} = p^-$ are the four-momenta of the initial quark and the antiquark constituting the system, respectively [11]. Although eq. (3.10) has been derived for a system in which q and $\overline{q}$ are moving in opposite directions, it is valid in all frames. Considering $s = E_{\mathrm{CM}}^2 = (p^+ + p^-)^2 = 2p^+p^- = \tilde{p}^+\tilde{p}^-$, where E_{CM} is the energy in the CM frame, the Lorentz invariant expression for hadron mass in terms of the $x^{\pm}$ fractions is

$$m_h^2 = p_h^2 = x_h^+ x_h^- s. \tag{3.11}$$

The initial system can also fragment generating a diquark pair instead of a $q\overline{q}$ pair. In such cases, the diquark combines with the quark from the adjacent breakup to create

a baryon. Similarly, the antidiquark combines with the adjacent antiquark generating another baryon. Apart from that, the procedure is the same as in meson production, both in the energy–momentum and the space–time pictures, meaning that all the properties presented for mesons in the fragmentation process also apply to baryons. Another way of generating baryons in the Lund String Model is through the "popcorn" model [19, 20].

Apart from the energy–momentum variables inherent to hadrons, each breakup has an associate parameter, Γ, which represents the squared invariant time of the breakup with respect to the origin [10, 11], scaled with the string tension

$$\Gamma = (\kappa\tau)^2 = \kappa^2 \left(t^2 - x^2 - y^2 - z^2\right). \tag{3.12}$$

The geometrical interpretation of Γ is depicted in Figure 4. Using the relations in eq. (3.8) and due to the linearity between energy–momentum and space–time pictures, the squared invariant time can be expressed in terms of the $\hat{x}^{\pm}$ variables [11] as

$$\Gamma = \left(\hat{x}^+ p^+ + \hat{x}^- p^-\right)^2 = \hat{x}^+ \hat{x}^- s. \tag{3.13}$$

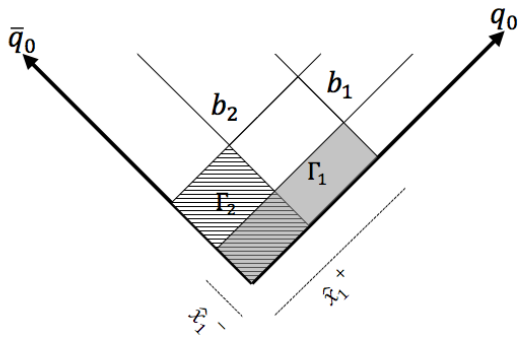


Figure 4: Squared invariant time representation of the breakups b_1 and b_2 in a $q\overline{q}$ system.

So far, the Lund String Model has been presented for a collinear $q\overline{q}$ system formed by massless quarks. This model can be easily generalized to more complicated systems, a topic that will be addressed in section 3.3. Before that, the generation of breakup points is first presented in the following section.

3.2.3 How are the breakup points obtained?

The fragmentation process is performed using the $z^{\pm}$ fractions, illustrated in Figure 3b, which relate the hadron four-momentum and the energy–momentum available in the system, $z^{\pm} = \tilde{p}_h^{\pm}/\tilde{p}_{\text{available}}^{\pm}$. It is important to emphasize here the difference between $z^{\pm}$ and the $x^{\pm}$ fractions, presented in the previous section. The latter correspond to the fractions taken by the hadron from the initial $\tilde{p}_{\text{total}}^{\pm}$ of the system, while the $z^{\pm}$ fractions are determined from the remaining $\tilde{p}^{\pm}$ after the creation of previous hadrons [10]. The difference between the two fractions is represented in Figure 3a and 3b. To illustrate the relation between

$x^{\pm}$ and $z^{\pm}$ fractions [10], we will consider a $q_0\overline{q}_0$ system, which fragments several times generating pairs $q_1\overline{q}_1$, $q_2\overline{q}_2$, etc. The first meson $q_0\overline{q}_1$ acquires a fraction $z_1^+ = x_1^+$ from the total $\tilde{p}^+$ of the system, while the remnant-jet is left with a $\tilde{p}^+$ fraction of $1 - z_1^+ = 1 - x_1^+$. The second hadron $q_1\overline{q}_2$ takes a fraction z_2^+ from the leftover $\tilde{p}^+$, i.e. $x_2^+ = z_2^+\left(1 - z_1^+\right)$. Thus, the remnant-jet carries the remaining fraction $\left(1 - z_1^+\right)\left(1 - z_2^+\right) = 1 - x_1^+ - x_2^+$. Since the fragmentation process is an iterative process, the $x^{\pm}$ fractions related to hadron i can be written as

$$\begin{aligned} x_i^+ &= z_i^+ \prod_{j=1}^{i-1}\left(1 - z_j^+\right) \\ x_i^- &= \frac{m_i^2}{x_i^+ s} \end{aligned} \tag{3.14}$$

where the relation for x_i^- is given in eq. (3.11). Although in this case we focus on the positive fractions, the same procedure can be followed for the negative fractions z^- and x^-. Since the same relations derived for z^+ can be derived for z^- following the equivalent procedure, we will refer to z^+ as z from now on.

One of the main properties of the fragmentation process is the fact that breakups are causally disconnected, which implies that the time ordering of the breaks is frame dependent. Even though in the centre of mass frame, the last hadrons are generated at the endpoints of the system, considering a process from the endpoints inwards is always more convenient to develop the fragmentation process. Although the breaks are causally disconnected, two adjacent string breaks have to be correlated such that the in-between hadron is created on its mass shell, since only the breakup points generating physical masses appear in the physical state [10].

So far, the fragmentation process has been developed as an iterative process starting from the q_0 side and moving towards the $\overline{q}_0$ end. Since the breakup points are causally disconnected, the same inverse procedure starting from the $\overline{q}_0$ side should result in the same set of hadrons on the average. From this property, known as left-right symmetry, and further considerations [21], the probability to create a hadron of mass m_h taking a fraction $z = z^+$ from the remaining four-momentum of the quark is given by

$$f(z) = \frac{(1-z)^a}{z}\exp\left(-b\frac{m_h^2}{z}\right) \tag{3.15}$$

where a and b are parameters that should be tuned to reproduce the experimental data. Hence, $f(z)$ determines how the individual vertices correlate in order to create a hadron of mass m_h by taking a fraction z of the energy–momentum left in the system.

Despite the acausal relation among breaks in the string, the squared invariant time of any breakup point can be expressed in terms of the previous breakup point as a pure geometrical relation

$$\Gamma_i = (1 - z_i)\left(\Gamma_{i-1} + \frac{m_h^2}{z_i}\right) \tag{3.16}$$

where m_h is again the mass of the hadron generated, z_i is the fraction of the energy–momentum taken by the hadron from the system, and Γ_i and Γ_{i-1} are the squared invariant times of the i and $i-1$ breakups, respectively.

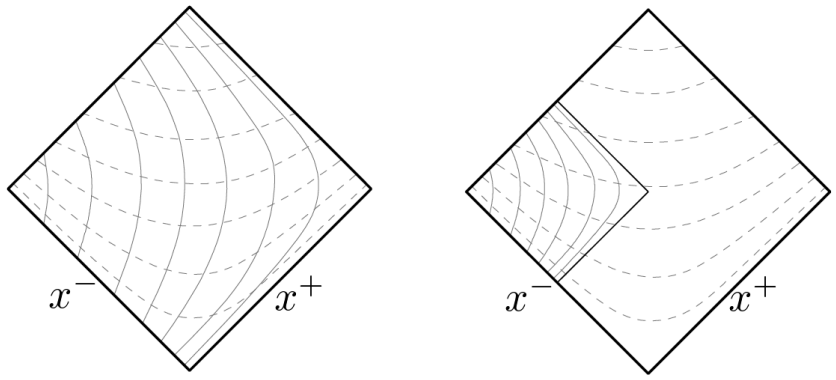
The inclusive probability for a breakup to take place at some specific Γ value is defined from the left-right symmetry [21] as

$$P(\Gamma) = \Gamma^a \exp\left(-b\Gamma\right) \tag{3.17}$$

with the same a and b as in eq. (3.15).

The breaks of the string can be determined by directly picking a z_i value, according to a given m_h value, by using eq. (3.15). Although this method works for simplest systems, such as $q\overline{q}$ systems, the z fractions cannot be easily extended to systems with more than two partons. Hence, another procedure is needed for the future that works in more complicated systems. This new procedure [10] uses the Lorentz invariant variable Γ_i and m_h^2, instead of z_i and m_h^2. First, a z_i value is picked which is later use to calculate Γ_i, from Γ_{i-1} and eq. (3.16). Then, if the breakup, along with the adjacent breakup, forms a hadron of mass m_h, its location should satisfy both eq. (3.11) and eq. (3.13).

This last procedure is graphically depicted for a $q\overline{q}$ system in Figure 5, where only one region needs to be considered. In that case, the Γ and m_h^2 distributions are hyperbolae, represented as dashed and full lines, respectively. The breakup location corresponds to a point where the hyperbolae for constant Γ and constant m_h^2 cross, since its location is determined by those two parameters. Figure 5a shows the hyperbolae related to the first closest breakup to the q end, while Figure 5b illustrates the corresponding setup for some random breakup inside the system. Due to left-right symmetry, the same procedure can be followed from the $\overline{q}$ end, in which case the m_h^2 hyperbolae are reverse with respect to the hyperbolae shown in Figure 5.



(a) First breakup from the endpoints (b) Some breakup inside the system

Figure 5: Hyperbolae of constant Γ and m_h^2 represented in dashed and full lines.

3.2.4 Tunneling process and transverse momentum p_T

Up to this point, we have assumed that the $q\overline{q}$ generated from string breaking are massless and have no transverse momentum, implying that both q and $\overline{q}$ are created as real particles at the same space–time location, with vanishing energy–momentum. If the pair is massive or carries transverse momentum, the transverse mass of q and $\overline{q}$ is determined as $m_T^2 = m^2 + p_T^2$, where m is the mass of the quark and p_T the transverse momentum. In such cases, the q and $\overline{q}$ are created in the same space–time location as virtual particles instead, and they tunnel out a distance $2l$ before turning into real particles. This distance $2l$ is such that the four-momentum stored in that string piece is used to give both q and $\overline{q}$ the corresponding transverse mass. Due to flavour and p_T conservation, q and $\overline{q}$ have the same mass and carry the same p_T but in opposite directions. Taking this into account, the distance $2l$ can be determined from the linear relation between space–time and energy–momentum as

$$2l\kappa = 2m_T \tag{3.18}$$

Therefore, pair production is a quantum mechanical process, which can then be considered a tunneling process, whose suppression factor is given by

$$\exp\left(-\frac{\pi m_T^2}{\kappa}\right) = \exp\left(-\frac{\pi m^2}{\kappa}\right)\exp\left(-\frac{\pi p_T^2}{\kappa}\right). \tag{3.19}$$

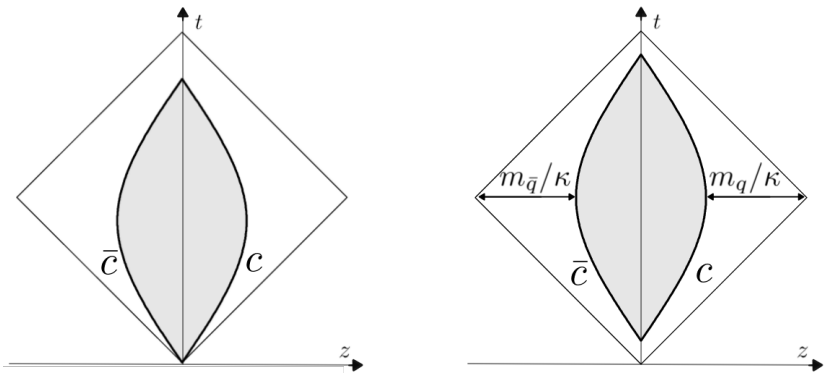
A consequence of this mechanism is the suppression of heavy quark production from string breaking, where pair production is estimated to be $u\overline{u} : d\overline{d} : s\overline{s} : c\overline{c} \approx 1 : 1 : \frac{1}{3} : 10^{-11}$ [10]. Since the production of heavy quarks, c and b quarks, is strongly suppressed, they are not generated as new $q\overline{q}$ pairs during the fragmentation process.

Taking into account that a meson is formed by the combination of a quark and an antiquark from two adjacent vertices, the transverse momentum of the meson is defined as the sum of the $\vec{p}_T$ vectors of the constituent quark and antiquark. As an example, the transverse momentum of meson $q_1\overline{q}_2$ is determined as $\vec{p}_T = \vec{p}_{T1} + \vec{p}_{T2}$, where $\vec{p}_{T1}$ and $\vec{p}_{T2}$ are the vectorial transverse momenta of q_1 and $\overline{q}_2$, respectively.

3.2.5 Massive quarks

Although massive quarks are not created from string breaking, they can be generated in the hard process and form a system that might fragment further. In this section, the yo-yo model is extended to account for massive quarks as the endpoints of the system. Since the massive q and $\overline{q}$ do not travel at the speed of light, their motion is described by hyperbolae instead of straight lines.

To study the motion of the massive yo-yo system, consider a $c\overline{c}$ system in the CM frame, in which c and $\overline{c}$ are moving along the z axis in opposite directions. The massive yo-yo system is depicted in Figure 6a, along with the massless case for comparison. At



(a) Motion of the $c\overline{c}$ system.

(b) Location of the $c\overline{c}$ system in the equivalent simple $q\overline{q}$ system.

Figure 6: The $c\overline{c}$ system and the equivalent system formed by massless q and $\overline{q}$. The grey region corresponds to the physical region.

time $t = 0$, the energy and the momenta of the massive quarks and the string are

$$\begin{aligned} E_c(0) = E_{\overline{c}}(0) = E_0 = \frac{E_{\mathrm{CM}}}{2} \\ p_{zc/\overline{c}}(0) = \pm p_0 = \pm\sqrt{E_0^2 - m_c^2} \\ E_{\mathrm{string}} = p_{z\mathrm{string}} = 0 \end{aligned} \tag{3.20}$$

where E_{CM} is the CM energy of the system. In the massive case, the linear relation between space–time and energy–momentum given by eq. (3.3) reduces to the proper relativistic definition of force, $dp_z/dt = \pm\kappa$, meaning that at time t

$$\begin{aligned} p_{zc/\overline{c}}(t) = \pm(p_0 - \kappa t) \\ E_{c/\overline{c}}(t) = E(t) = \sqrt{p_{zc/\overline{c}}^2(t) + m_c^2}. \end{aligned} \tag{3.21}$$

From the energy expression, the speed at which c and $\overline{c}$ travelled can be determined, i.e $dE(t)/dt = -\kappa p_{zc/\overline{c}}(t)/E(t) = -\kappa v_{zc/\overline{c}}(t)$, with $v_{zc}(t) = -v_{z\overline{c}}(t)$. The distance travelled by c or $\overline{c}$ at time t is

$$z_{c/\overline{c}}(t) = \int_0^t v_{zc/\overline{c}}(t)\,dt = \pm\frac{1}{\kappa}(E_0 - E(t)), \text{ so } E_{c/\overline{c}}(t) = E(t) = E_0 \mp \kappa z_{c/\overline{c}}(t). \tag{3.22}$$

Note that $z_c(t) = -z_{\overline{c}}(t) = z(t)$, since in the CM of mass frame $E_c(t) = E_{\overline{c}}(t)$. Hence, the four-momenta of c, $\overline{c}$ and the string at time t can be expressed in a similar way as in the the massless yo-yo case

$$\begin{aligned} p_{c/\overline{c}}(t) = (E_0 \mp \kappa z(t);\ 0,\ 0, \pm(p_0 - \kappa t)) \\ p_{\mathrm{string}}(t) = (2\kappa z(t);\ 0,\ 0,\ 0) \end{aligned} \tag{3.23}$$

which implies that the motion of the massive system can be explained similarly to the procedure presented in section 3.2.1. Although the motion properties of the massless

and massive cases hold in every longitudinal boosted frame due to Lorentz covariance, the boosted system is simpler to address when considering massless quarks. Moreover, the study of the kinematics is easier in the massless case, since for massless four-vectors $p_q^2 = p_{\overline{q}}^2 = 0$, a property that does not hold for massive systems. It is for these reasons that the kinematics of the massive case is developed in terms of an equivalent system formed by massless quarks, whose string has the same properties as the string from the massive case.

In order to study the kinematics of the massive case, the trick is to replace the quark masses by a string piece at each endpoint. The length of the extra piece at each end is given by $l = m_q/\kappa$ at the turning point, when $p_z = 0$. This situation is depicted in Figure 6b, where the massive motion is illustrated by the hyperbolae, whose asymptotes are the straight lines of the massless case. Since the energy stored by this extra string piece is used to give mass to the initial quarks, it does not break during the fragmentation process. The region that can fragment is known as physical region and corresponds to the region between the hyperbolae, highlighted in grey in Figure 6b. The first hadron created from the endpoint is always heavier than the massive quark, $m_{\text{had}} > m_q$ and, consequently, all the hadrons are automatically created inside the physical region.

To define the equivalent massless case, the massless reference four-vectors should be determined. Due to Lorentz covariance, the q and $\overline{q}$ four-momenta of the massless case have to be linear combinations of the massive quark four-momenta. If p_q and $p_{\overline{q}}$ are the four-momenta of the massive quarks, while $\tilde{p}_q$ and $\tilde{p}_{\overline{q}}$ the massless four-momenta, the relation between both systems is given by

$$\begin{aligned} \tilde{p}_q &= (1 + k_1)\, p_q - k_2 p_{\overline{q}} \\ \tilde{p}_{\overline{q}} &= (1 + k_2)\, p_{\overline{q}} - k_1 p_q \end{aligned} \tag{3.24}$$

where the expressions for k_1 and k_2 are derived using the above relations and the fact that $\tilde{p}_q^2 = \tilde{p}_{\overline{q}}^2 = 0$.

The massless and massive systems are aligned so that the turning points of the partons occur at the same time, with m_q/κ being the separation between the endpoints of massive and massless systems. Consequently, the point in which the massive oscillation begins does not correspond to the real origin of the system, an offset that has to be taken into account when defining the space–time location of breakups (see section 4.5), since the fragmentation process takes place in the equivalent massless system.

3.3 Multiparton systems

At the LHC, much more complicated processes than a pure $q\overline{q}$ system take place. An example of such processes is the Z^0 decay into a pair of massless quarks, which can emit a gluon:

$$Z^0 \longrightarrow q\overline{q} \longrightarrow qg\overline{q}.$$

The gluon emission in the Z^0 decay is part of the partonic process, where the interaction among partons occur, just after high-energy particles collide. In such processes, gluon radiation is generated from virtual quarks, whose lifetimes are normally short, meaning that they do not travel far away from the creation point before emitting a gluon. These processes are known as parton showers and take place before the hadronization process. In an LHC event, where the hadronization process can extend over several thousand fermi, the parton showers in one part of the event might take place after the hadronization process has started in another part. Nevertheless, locally, parton showers occur some time before the hadronization process. Therefore, the gluons emitted from gluon radiation are considered to be created at the same space–time location as the initial quark and antiquark, if the gluons are energetic and well separated from the rest of the partons forming the system.

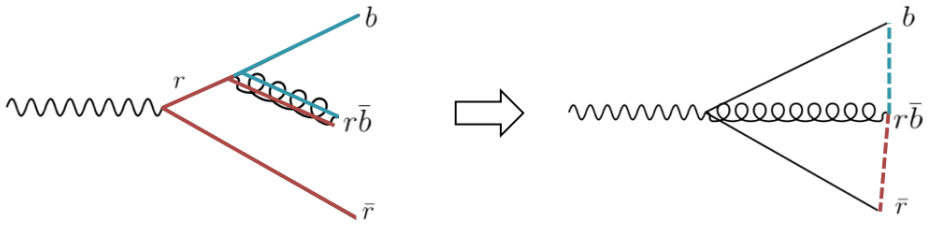


Figure 7: Colour flow (left) in the Z^0 decay system and approximation in the Lund String Model (right).

Another important property to deal with when accounting for gluons is colour flow. In order to understand how this property is implemented in the Lund String Model, we first consider that a non-diagonal gluon is emitted in the Z^0 decay. Such gluons carry separate colour and anticolour [1], implying that the colour flows in the system as exemplified in the left diagram of Figure 7. In this case, the colour flow is modelled in the Lund String Model by the existence of two string regions, between q and g and between g and $\overline{q}$, as illustrated in the right diagram of Figure 7. The same procedure cannot be followed when the gluon emitted is a diagonal one, such as $\left(r\overline{r} + g\overline{g} - 2b\overline{b}\right)/\sqrt{6}$, because colour assignment is not obvious anymore. Nevertheless, the fraction of diagonal gluons decreases with increasing number of colours in an SU(N) theory, making the number of diagonal gluons negligible in a theory with an infinity number of colours. This limit is assumed in the Lund String Model, since it gives rise to an unambiguous colour assignment.

Once those assumptions are taken into account, the string dynamics of the $qg\overline{q}$ system formed by massless partons can be addressed. By adopting a suitable Lorentz frame, the $qg\overline{q}$ system can always be studied in a frame in which q moves in the $+z$ direction, g in the $+x$ direction and $\overline{q}$ in the $-z$ direction. As explained before, two string pieces are present at the beginning of the process, the string connecting q and g and the string between g and $\overline{q}$, as illustrated in diagram 1 of Figure 8. Each string piece is regarded as a separate string region, which behaves similarly to the string piece of a $q\overline{q}$ system. To study the string dynamics, we will focus on the first region formed by a string between q and g. The

four-momenta of q, g and the string at $t = 0$ are

$$\begin{aligned} p_q\,(0) &= E_q\,(1;\ 0,\ 0,\ 1) \\ p_g\,(0) &= E_g\,(1;\ 1,\ 0,\ 0) \\ p_{\text{string}}\,(0) &= (0;\ 0,\ 0,\ 0)\,. \end{aligned} \tag{3.25}$$

Following a similar approach than for a $q\overline{q}$ system, the system evolves with time as

$$\begin{aligned} p_q\,(t) &= (E_q - \kappa t)\,(1;\ 0,\ 0,\ 1) \\ p_g\,(t) &= (E_g - 2\kappa t)\,(1;\ 1,\ 0,\ 0) \\ p_{\text{string}}\,(t) &= \kappa t\,(2;\ 1,\ 0,\ 1)\,. \end{aligned} \tag{3.26}$$

The behaviour of the string connecting g and $\overline{q}$ can be derived exactly in the same way, taking into account that $\overline{q}$ moves along the $-z$ direction. Note the factor of two in the gluon four-momentum, which stands for the presence of the two strings connecting g with q and $\overline{q}$, respectively. Contrary to the $q\overline{q}$ system, the two string pieces are not at rest, but move in the transverse direction: the qg string piece has a transverse motion $v_x = v_z = \frac{1}{2}$, while the $g\overline{q}$ piece $v_x = -v_z = \frac{1}{2}$. Although the energy per unit string length is higher than for a string at transverse rest, the force acting on the endpoints is of the same magnitude. Consequently, the same κ value for the $q\overline{q}$ system can also be used for more complicated topologies [11].

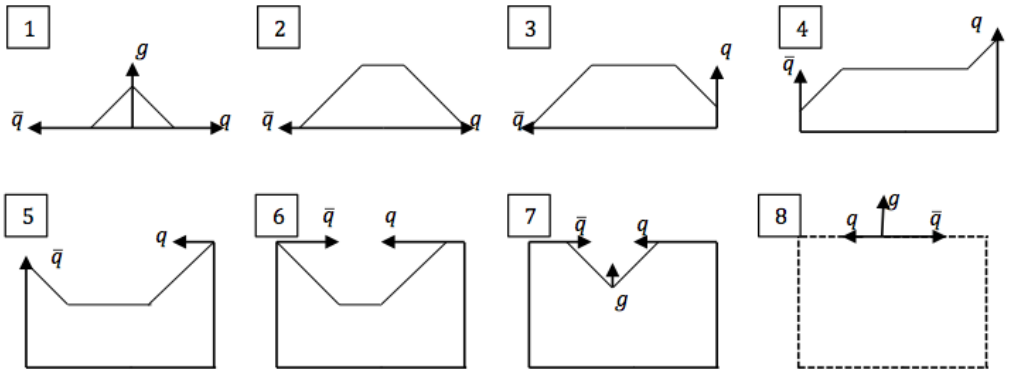


Figure 8: Time evolution of the $qg\overline{q}$ system formed by massless partons in a frame where the gluon moves in the $+x$ direction, while the q and $\overline{q}$ move in opposite directions along the z axis. Figure adapted from [11].

Figure 8 illustrates the evolution of half a period of the $qg\overline{q}$ system in the absence of string breaks. As indicated before, the system is initially formed by two string regions: the qg and the $g\overline{q}$ regions, as depicted in diagram 1. After some time, we assume that the gluon loses all its energy and a new region is created, which we refer to as the $q\overline{q}$ region, since the string piece takes energy and momentum from the q and $\overline{q}$, see diagram 2. At some point, the q also transfers all its energy to the string and starts moving in the $+x$ axis, due to the influence of the string, which transfers the momentum taken from the gluon to

the quark. The q eventually gains half of the gluon energy, which is then given back to the string again, as illustrated in diagrams 3 and 4 from Figure 8. Therefore, the quark finally loses its energy and starts moving along the $-z$ direction, due to the string effect. A similar process occurs for $\overline{q}$. As shown in diagram 7 of Figure 8, the gluon will eventually appear again and the sequence will be repeated with the momenta of q and $\overline{q}$ swapped. If breaks are taken into account, the evolution of the system is much more complicated.

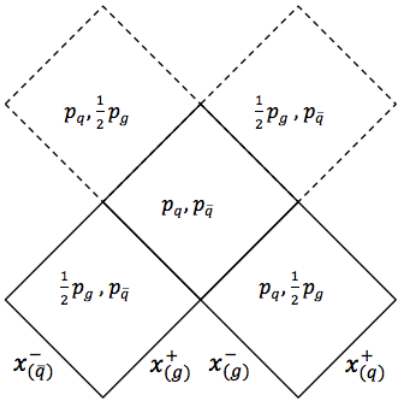


Figure 9: The parameter plane picture for the $qg\overline{q}$ system. The dash lines indicate the turnover regions, normally neglected. Figure adapted from [11].

Although the diagrams from Figure 8 are useful to analyse the time evolution of the system, the parameter plane picture is most convenient when addressing the kinematics. The parameter plane picture [11] is a diagram that displays the different string regions in terms of the light-cone four-vectors defining each region, i.e. $p_q^+ = p_q$, $p_{\overline{q}}^- = p_{\overline{q}}$ and $p_g^+ = p_g^- = \frac{1}{2}p_g$ in the $qg\overline{q}$ case. Figure 9 depicts the parameter plane for the $qg\overline{q}$ system discussed so far. The low regions represent the states in which none of the partons have lost their energy, corresponding to the two string regions in diagram 1 of Figure 8, the qg and the $g\overline{q}$ string pieces. The intermediate region corresponds to the new string piece created from the q and $\overline{q}$ momenta once the gluon has lost all its energy, which is represented in diagram 2 of Figure 8. Finally, the upper regions account for the turnover regions, which are related to the two string pieces formed when g re-appears and q, $\overline{q}$ swap their momenta, both illustrated in diagram 8 of Figure 8. Although the complete parameter plane picture is the one shown in Figure 9, the upper regions or turnover regions are normally neglected, since we assume the system hadronizes before the q and the $\overline{q}$ swap their momenta. This reasonable assumption is made due to large number of complications in the fragmentation pattern included when accounting for the turnover regions [11].

From the parameter plane picture, the equations defining the hadron properties and the fragmentation process presented in section 3.2 can be easily generalized. For the $qg\overline{q}$ system, the most generatic form of expressing the hadron four-momentum in terms of the $x^{\pm}$ fractions is

$$\begin{aligned} p_h &= x_q^+ p_q^+ + x_g^+ p_g^+ + x_g^- p_g^- + x_{\overline{q}}^- p_{\overline{q}}^- \\ &= x_q^+ p_q + \frac{1}{2}\left(x_g^+ + x_g^-\right) p_g + x_{\overline{q}}^- p_{\overline{q}}. \end{aligned} \tag{3.27}$$

The hadron mass can be determined directly from the four-momentum as $m_h^2 = p_h^2$. The other Lorentz invariant variable Γ is obtained from the $\hat{x}^{\pm}$ defined in section 3.2.2 as

$$\Gamma = \left(\hat{x}_q^+ p_q^+ + \hat{x}_g^+ p_g^+ + \hat{x}_g^- p_g^- + \hat{x}_{\overline{q}}^- p_{\overline{q}}^-\right)^2 . \tag{3.28}$$

Note that the relation between the hadron momentum and the light-cone coordinates $\tilde{p}^{\pm}$, introduced for the single string system, is more complicated in this case. In a frame where g moves along the $+x$ direction, q along the $+z$ and $\overline{q}$ along the $-z$ direction, the light-cone coordinates of the hadron are [11]

$$\begin{aligned} \tilde{p}_h^+ &= E + p_z = 2x_q^+ E_q + \frac{1}{2}\left(x_g^+ + x_g^-\right) E_g \\ \tilde{p}_h^- &= E - p_z = 2x_{\overline{q}}^- E_{\overline{q}} + \frac{1}{2}\left(x_g^+ + x_g^-\right) E_g \\ p_x &= \frac{1}{2}\left(x_g^+ + x_g^-\right) E_g \end{aligned} \tag{3.29}$$

where E_q is the energy of the massless quark, E_g the energy of the gluon, $E_{\overline{q}}$ the energy of the massless antiquark and $x_g^{\pm}$, $x_q^{\pm}$ and $x_{\overline{q}}^{\pm}$ are the hadron $x^{\pm}$ fractions related to the gluon, the quark and the antiquark, respectively. Therefore, the hadron will have a transverse momentum and a transverse mass-square, defined as $m_T^2 = (E + p_z)(E - p_z) = \tilde{p}_h^+ \tilde{p}_h^-$, which does not match the total hadron, determined as $m_h^2 = m_T^2 - p_x^2$.

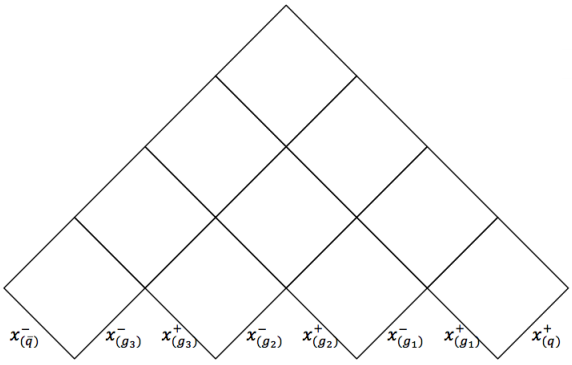


Figure 10: The parameter plane picture for a multiparton system composed by five partons.

The parameter plane picture can be easily extended to a multiparton system, resulting in the most convenient alternative to study the kinematics of any multiparton system. As an example, the parameter plane for a system consisting of three gluons, one quark and one antiquark is depicted in Figure 10, where the turnover regions have been ignored. In such system, there are four low regions, or initial regions, and six intermediate regions. The number of initial and intermediate regions can be generalized for any multiparton system formed by n partons, out of which $n-2$ are gluons, as $n-1$ initial regions and $\frac{1}{2}(n-1)(n-2)$ intermediate regions [11]. The expression for the hadron four-momentum can be also conveniently generalized to an n-parton system by accounting for the momenta

taken from each parton as

$$p_h = x_q^+ p_q^+ + x_{\overline{q}}^- p_{\overline{q}}^- + \sum_{i=1}^{n-2} \left(x_{g_i}^+ p_{g_i}^+ + x_{g_i}^- p_{g_i}^- \right). \tag{3.30}$$

Apart from these aspects, the rest of the properties are similarly determined as in previous cases for a single string system or $qg\overline{q}$ systems.

3.4 Lund String Model implementation

The hadronization process in PYTHIA starts once the initial partonic system has been built, based on the four-momenta of the initial partons, previously created during the hard process. Due to left-right symmetry, PYTHIA first chooses randomly the side from which the step is taken to create a new $q\overline{q}$ pair, either from the q or the $\overline{q}$ end. Then, the flavour of the $q\overline{q}$ pair is picked randomly, taking into account that the production of $s\overline{s}$ is suppressed as $u\overline{u} : d\overline{d} : s\overline{s} \approx 1 : 1 : \frac{1}{3}$ [10]. As mentioned before, heavy quarks are not generated during the fragmentation process due to the strong suppression given by the tunneling factor (see section 3.2.5).

Once the flavour of the new $q\overline{q}$ is set, a hadron is formed combining the old q with the new $\overline{q}$, if the step has been taken from the q end, and vice versa otherwise. The hadron kind is then established depending on the quark content. Since different hadrons share the same quark content, this decision is basically based in two different effects with opposite bias: the spin effect and the mass effect. Due to the spin effect, particle with higher spin number are generated more often than lower spin particles, because the number of states goes like $2s+1$, where s is the spin number. On the other hand, more massive hadrons are less probable to be generated than lighter hadrons. Because their bias are inverse, those two effects balanced when the hadron kind is established.

After the hadron kind is established, the hadron mass is set and the transverse momentum is summed from the constituent q and $\overline{q}$ contributions. Finally, the longitudinal momentum is determined and added to the total hadron four-momentum. In order to determined the longitudinal momentum, a z_i value is randomly picked according to eq. (3.15), which, in a simple system, corresponds to z_i^+ if the step has been taken from the q end and to z_i^- otherwise. For that specific z_i value, the Γ_i value of the breakup is calculated using eq. (3.16). In a simple system, the breakup space–time location is obtained from the m_h^2 and the z_i value, while in a multiparton topology, a region is searched that obeys the $\{m_h^2, \Gamma_i\}$ constraint, which gives the final solution for the breakup location in the energy–momentum picture. The hadron longitudinal momentum is then determined as stated in section 3.2.2. This process is repeated until the energy left in the system is not enough to produce three hadrons more. At that point, a final breakup is generated in between the two last breakups from each side, creating two final hadrons. Details on the generation of the final two hadrons can be found in Appendix A, along with the challenges encountered when implementing the space–time picture and the methods applied to solved them.

4 Space–time implementation

4.1 Simplest $q\overline{q}$ system and hadron production points

So far, the hadronization process in PYTHIA was developed in terms of the energy–momentum fractions $x^{\pm}$ and $z^{\pm}$, presented in section 3.2.2. Therefore, the location of the breakup points could only be determined in the energy–momentum picture, in terms of the $(\tilde{p}^{+}, \tilde{p}^{-})$ coordinates. In order to study the density of hadron production, the energy–momentum picture should first be translated to the space–time picture, defined in terms of (t, x, y, z) coordinates.

To illustrate the space–time implementation in the hadronization process, we first consider a breakup point i in a $q\overline{q}$ system. Its location with respect to the origin of the energy–momentum picture, where q and $\overline{q}$ have been created, is given by the $\hat{x}^{\pm}$ fractions. Then, considering p^{+} to be the q four-momentum and p^{-} the $\overline{q}$ four-momentum, the location of breakup i in the energy–momentum picture is defined as $\hat{x}_i^{+}p^{+} + \hat{x}_i^{-}p^{-}$. Recalling the linear relation between space–time and energy–momentum (eq. (3.3)), the space–time location of breakup point i is defined as

$$v_i = \frac{\hat{x}_i^{+}p^{+} + \hat{x}_i^{-}p^{-}}{\kappa}. \tag{4.31}$$

Note that the q and $\overline{q}$ generated by string breaking are considered to be created in the same space–time location, even if the transverse momentum and the masses of the quarks are considered. Although Heisenberg uncertainty relations hold and tunneling is taken into account in the Lund String Model, this is the best assumption we can make in order to study the space–time behaviour of hadron production.

Eq. (4.31) determines the space–time location of the string breakups inside any region. Because hadrons are formed by two adjacent string breakups, the space–time location at which they are generated can be determined from the location of the constituent breakups. Contrarily to the space–time location of a breakup, the hadron production point definition is not unique, since hadrons are composite particles. For that reason, we propose three alternative definitions of hadron production points, illustrated in Figure 11. Figure 11 displays a $q_0\overline{q}_0$ system with two breakup points, i and $i+1$, whose space–time coordinates are v_i and v_{i+1}, respectively. The $q_i\overline{q}_{i+1}$ subsystem forms hadron i, whose hadron production point, v_i^h, can be defined by the middle point between the two breakups i and $i+1$, red dot in Figure 11, as

$$v_i^h = \frac{v_i + v_{i+1}}{2}. \tag{4.32}$$

Alternatively to this "middle" definition, the "late" hadron production point is interpreted as the point where the two partons forming the hadron cross for the first time, illustrated in green in Figure 11. Taking into account the hadron four-momentum, p_h, the "late" hadron production point is determined from the "middle" definition as

$$v_i^h = \frac{v_i + v_{i+1}}{2} + \frac{p_h}{2\kappa}. \tag{4.33}$$

Finally, an "early" description, represented in blue in Figure 11, is defined as the initial point of the fictional oscillation between the partons that form the hadron, given by

$$v_i^h = \frac{v_i + v_{i+1}}{2} - \frac{p_h}{2\kappa}. \tag{4.34}$$

Note that the two endpoint hadrons are situated on the light-cone with this "early" definition. The different results drawn from the three alternative definitions can be used as a measure of uncertainty, as it is shown in section 5.1. If not stated otherwise, the default value use in this thesis is the "middle" definition, given by eq. (4.32).

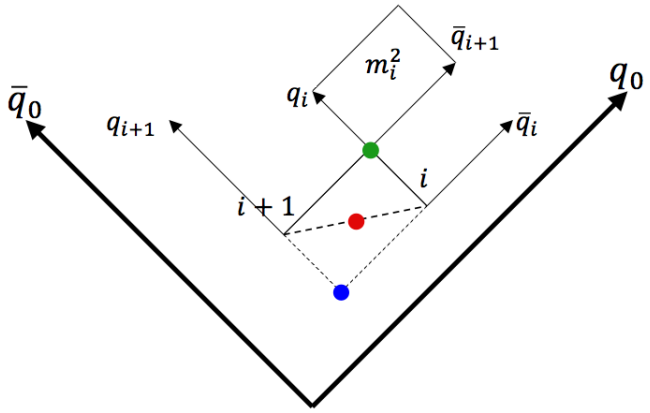


Figure 11: Hadron formation in a $q\bar{q}$ system. The blue, red and green dots represent the "early", "middle" and "late" definitions of hadron production points, respectively.

4.2 More complex topologies

Multiparton systems are more complicated to address than single string systems, as shown for the string dynamics and the kinematics in section 3.3. Their complexity also affects the space–time implementation, which has to be extended to include several string regions. To explain how the space–time picture is generalized to multiparton systems, consider a $qg\bar{q}$ system formed by massless partons, whose parameter plane is displayed in Figure 12, ignoring the turnover regions. Since each string region behaves as a simple $q\bar{q}$ system, eq. (4.31) can be used within each region. Nevertheless, the intermediate region is formed after the gluon has lost all its energy, at a different location in space–time than the initial regions, an offset that has to be taken into account when determining the breakup location in space–time. From the linear relation between space–time and energy–momentum, the offset for this region can be calculated as $v_{\text{reg}} = p_g/2\kappa$, where p_g is the four-momentum of the gluon. The factor of $1/2$ in the offset accounts for the fact that the gluon loses four-momentum twice as fast as the q or $\bar{q}$, since it transfers its four-momentum to two string pieces. Then, the space–time location of a breakup located in the intermediate region is given by

$$v_i = \frac{\hat{x}_i^+ p^+ + \hat{x}_i^- p^-}{\kappa} + \frac{p_g}{2\kappa}. \tag{4.35}$$

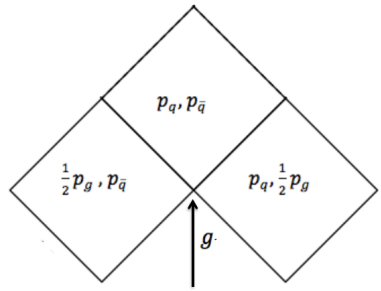


Figure 12: Parameter plane for a $qg\overline{q}$ system.

If the system is composed of more than one gluon, more than one intermediate region have to be taken into account, as illustrated in Figure 13 for system composed of three gluons, one quark and one antiquark. In such cases, more gluons have to be included when determining the space–time offset of some intermediate regions, such as region qg_3 in Figure 13. The region qg_3 is created when the gluons g_1 and g_2 have lost their energies, implying that the region offset in space–time is $v_{\text{reg}} = (p_{g_1} + p_{g_2})/2\kappa$. Following a similar procedure, the offset of any region in a multiparton system can be determined, which has to be added to the space–time location of the breakups situated in that region.

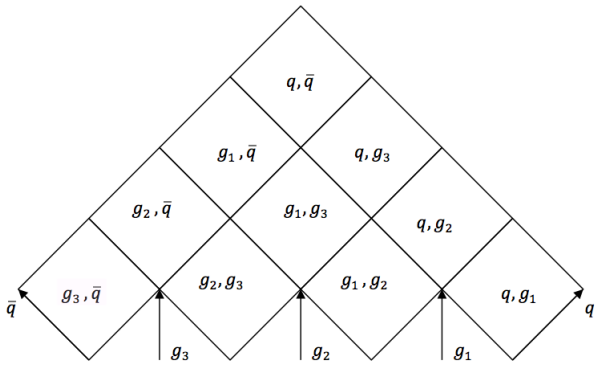


Figure 13: Parameter plane for a five parton system.

A general expression for the space–time offset of any intermediate region in any multiparton system can be easily defined. For a system composed of n partons, out of which $n-2$ are gluons, assign each parton a number such that q is $\tilde{p}_0$, $\overline{q}$ is $\tilde{p}_{n-1}$, g_1 is $\tilde{p}_1$, etc. With this notation, the offset of the region $\tilde{p}_j\tilde{p}_k$ is found to be

$$v_{\tilde{p}_j\tilde{p}_k} = \sum_{m=j+1}^{k-1} \frac{p_m}{2\kappa} \tag{4.36}$$

where p_m is the four-momentum of parton m. Thus, the general expression of the space–

time location for a breakup point i located in region $\tilde{p}_j\tilde{p}_k$ is

$$v_i = \frac{\hat{x}_i^+ p_j^+ + \hat{x}_i^- p_k^-}{\kappa} + \sum_{m=j+1}^{k-1} \frac{p_m}{2\kappa}. \tag{4.37}$$

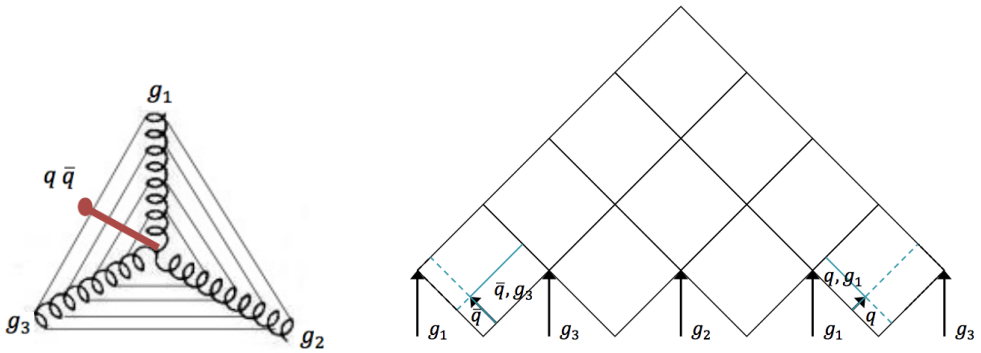
The implementation of the fully space–time picture in multiparton systems has been a great challenge in this thesis due to several complications found when the fragmentation process jumps among different regions. This aspect gives rise to two main problems: the determination of the space–time location of the final breakup in the system and the non physical values of the $\hat{x}^\pm$ fraction when fragmentation moves to a new region. Those issues are further explained in appendices A and B, respectively, along with the solutions found to properly implement the space–time picture.

4.3 Gluon–loops

So far, we have just taken into account gluons when constituting multiparton systems, along with a quark and an antiquark. However, gluons are the force carries of QCD and, since they carry both colour and anticolour, they both interact with quarks and other gluons. The latter interaction can lead to closed gluon–loop topologies, whose string configurations are more complex to address than multiparton systems, since the strings form a closed string system, as exemplified in Figure 14a for a ggg system. As a consequence, the relation of the string system with the parameter plane is not so trivial anymore. In this section, we explain how to define a parameter plane for a gluon-loop topology by breaking one string piece into two pieces, by the generation of a $q\overline{q}$ pair.

In order to reduce the problem to a familiar one, a $q\overline{q}$ is generated by string breaking in one of the string regions, fragmenting that region into two regions with smaller invariant mass. First, the region in which the pair is created is randomly selected, provided that the probability to generate at least one break in that region is favoured. Note that, when choosing the region, a bias is included in the fragmentation process. Once the region is selected, the breakup point is randomly allocated in the energy–momentum picture, always obeying the expected Γ distribution, eq. (3.17). The process followed to determine the energy–momentum location of the new $q\overline{q}$ pair in the specific region is similar to the one followed by the fragmentation process presented in section 3.4. After this process, the gluon-loop topology can be represented by a two-dimensional parameter plane and treated as a multiparton system, with the only difference that the q and $\overline{q}$ endpoints are not different than the rest of breakups, since $\Gamma_q = \Gamma_{\overline{q}} \neq 0$.

As an example, the parameter plane for a gluon–loop consisting of three gluons is displayed in Figure 14b. In this case, the string between g_1 and g_3 has broken into two string pieces, generating two new string regions: one formed by the influx of g_1 and q momenta and the other by g_3 and $\overline{q}$ momenta. The g_1g_3 region is duplicated in the parameter plane, such that the right endpoint region makes use of the "active area" between q and g_1, while the other uses the complementary area between g_3 and $\overline{q}$. Note that, since the q and $\overline{q}$



(a) String configuration (b) Parameter plane

Figure 14: The string configuration and the corresponding parameter plane for a three gluon-loop topology.

endpoints are generated from string breaking as a new $q\overline{q}$ pair, they are created at the same location in space–time. Apart from that, the fragmentation process is carried out in the same way as for any multiparton system and, hence the space–time location of the breakups can be determined as described in sections 4.1 and 4.2.

4.4 Smearing in transverse space

The hadronization process is simulated in Pythia by assuming the existence of a zero-width string connecting partons. As a better approximation, the QCD potential can be expressed as a string whose width is not zero, in which case the space–time location of the breakups has a transverse uncertainty with respect to the string axis. In this section, the calculation of the transverse space components is presented. This smearing in transverse space is included after the longitudinal space–time location has been derived from the fragmentation scheme.

The transverse location of breakup points is assumed to follow a Gaussian distribution, whose expectation value is zero. Thus, the probability density for the three-dimensional transverse space–time location of the breakups is

$$f\left(x,y\right) \propto e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{4.38}$$

where x and y are transverse spatial coordinates and σ is the standard deviation of the distribution. Because we are dealing with hadron formation, the string width can be approximated to be the proton size, whose three-dimensional radius is $r_p \approx 0.87$ fm [22]. Then, the standard deviation in each direction, σ, corresponds to

$$r_p^2 = \langle x^2 + y^2 + z^2 \rangle = 3\sigma^2, \text{ which implies } \sigma = \frac{r_p}{\sqrt{3}}. \tag{4.39}$$

Since the smearing in transverse space only accounts for two dimensions in space, the

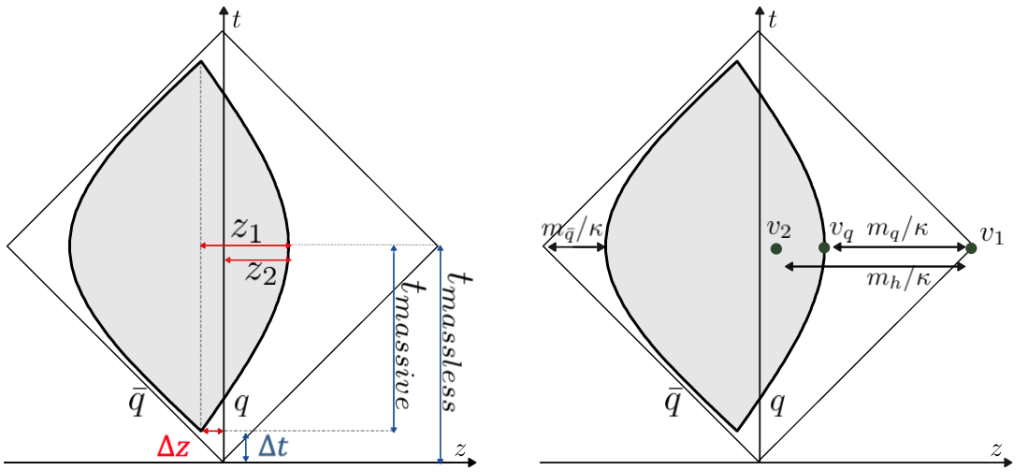
transverse radius of the string is determined as

$$r_{pT}^2 = \langle x^2 + y^2 \rangle = 2\sigma^2 = \frac{2}{3} r_p^2. \tag{4.40}$$

The smearing in transverse space might generate unwanted situations, such as negative values for the Γ parameter of the breakup points. Since the space–time location is first obtained from the fragmentation picture in the longitudinal direction, the squared invariant time should not change when introducing smearing. Therefore, the time coordinate is adjusted after including the smearing in transverse space, in order to retain the Γ value determined by the longitudinal scheme.

4.5 Massive quark implementation

As illustrated in section 3.2.5, the origin of the massive and massless oscillations are displaced for technical reasons, correspondingly the intial point of the massive oscillation is offset from the origin of the space–time coordinate system. Since the fragmentation process is performed in the massless system, the space–time locations of the breakups have to be adjusted to lead to the correct locations.



(a) Massive offset calculation

(b) Endpoint correction

Figure 15: Massive $q\overline{q}$ system and equivalent massless system. The grey region corresponds to the physical region.

To determine the deviation of the massive system form the origin, consider the system from Figure 15a, which is composed of two massive q and $\overline{q}$ in the CM frame moving in the $+z$ and $-z$ direction, respectively. In this system, the masses of q and $\overline{q}$ are different, with $m_q > m_{\overline{q}}$. Note that, at time $t = 0$, the system is in the centre of mass frame, where

$p_0 = p_{zq} = -p_{z\overline{q}}$ and $E_{\rm CM} = E_q + E_{\overline{q}}$, and the energy and momentum of q and $\overline{q}$ are

$$\begin{aligned} E_q &= \frac{E_{\rm CM}^2 + m_q^2 - m_{\overline{q}}^2}{2E_{\rm CM}} \\ E_{\overline{q}} &= \frac{E_{\rm CM}^2 + m_{\overline{q}}^2 - m_q^2}{2E_{\rm CM}} \\ p_0 &= \frac{\sqrt{\left(E_{\rm CM}^2 - m_q^2 - m_{\overline{q}}^2\right)^2 - 4m_q^2 m_{\overline{q}}^2}}{2E_{\rm CM}}. \end{aligned} \tag{4.41}$$

The massive oscillation in Figure 15a is offset both in time and z-component of space, represented as Δt and Δz. The temporal offset is determined from the difference between the time coordinates at which the massless and massive quarks lose their three–momenta, $t_{\rm massless}$ and $t_{\rm massive}$ in Figure 15a, respectively, whose values are given by $t_{\rm massless} = E_{\rm CM}/2\kappa$ and $t_{\rm massive} = p_0/\kappa$. Consequently, the time offset of the massive oscillation is

$$\Delta t = t_{\rm massless} - t_{\rm massive} = \frac{E_{\rm CM} - 2p_0}{2\kappa}. \tag{4.42}$$

The process to define the space offset is slightly different. In Figure 15a, the distance of the massive q endpoint to the centre of the massive oscillation is represented as z_1, while z_2 corresponds to the distance of the massive q endpoint to the centre of the massless system. These two quantities can be determined using eq. (3.22) and analysing the motion of the massless system. As illustrated in Figure 15a, the z-offset is given by the difference between z_1 and z_2

$$\Delta z = z_1 - z_2 = \frac{1}{\kappa}\left(\frac{E_{\rm CM}}{2} - m_q\right) - \frac{1}{\kappa}\left(E_q - m_q\right) = \frac{E_{\overline{q}} - E_q}{2\kappa}. \tag{4.43}$$

The time and space offsets, given by eq. (4.42) and eq. (4.43), can be written in terms of the light-cone coordinates as

$$\begin{aligned} \Delta t - \Delta z &= \frac{E_q - p_0}{\kappa} \\ \Delta t + \Delta z &= \frac{E_{\overline{q}} - p_0}{\kappa}. \end{aligned} \tag{4.44}$$

To define a general expression for the offset in space–time, the expressions in eq. (4.44) are related to the vectors of the massless system, $p^{\pm}$, given by the four-momenta of the massless quarks. Hence, the vectorial form of the space–time massive offset is

$$v_{\rm offset} = \frac{1}{\kappa E_{\rm CM}}\left(\left(E_{\overline{q}} - p_0\right)p^+ + \left(E_q - p_0\right)p^-\right). \tag{4.45}$$

Hence, for each vertex in a region formed by at least one massive quark, the space–time location is defined as usual from the massless system, $v_{\rm massless}$, and then corrected to obtain the correct space–time location

$$v_{\rm correct} = v_{\rm massless} - v_{\rm offset}. \tag{4.46}$$

For more complex topologies, such as multiparton systems consisting of a massive q and/or $\overline{q}$ and several gluons, the effect of the massive q or $\overline{q}$ is only non-negligible in the lowest respective endpoint region. Therefore, the correction due to the presence of massive quarks is only performed in those regions.

Whenever a system is formed by massive quarks, the space–time locations of the endpoints also have to be corrected, because the endpoints obtained from the fragmentation process correspond to the endpoints of the massless system. This situation is exemplified in Figure 15b, which illustrates the same massive $q\overline{q}$ system as before. The three vertices v_1, v_q and v_2 sketched correspond to the space–time location of the massless endpoint, the massive turning point and the closest breakup to the massless endpoint, respectively. Since the motion takes place along the z axis, the x and y components of the vertices are equal to zero. The system can be studied in a Lorentz frame such that the three vertices are simultaneous in time, $v_{1t} = v_{2t} = v_{qt}$. In such frame, the linearity between energy–momentum and space–time (eq. (3.3)) gives rise to the relations

$$\begin{aligned} v_{1z} - v_{2z} &= \frac{m_h}{\kappa} \\ v_{1z} - v_{qz} &= \frac{m_q}{\kappa} \end{aligned} \tag{4.47}$$

where m_q is the mass of the heavy quark and m_h the mass of the hadron formed from the vertices v_1 and v_2. Then, v_{zq} can be determined as $v_{qz} = v_{1z} + \frac{m_q}{m_h}(v_{2z} - v_{1z})$. Using four-vector notation, the general expression of the space–time location of the massive endpoint is given by

$$v_q = v_1 + \frac{m_q}{m_h}(v_2 - v_1)\,. \tag{4.48}$$

Eq. (4.48) is Lorentz invariant and can be applied to both q and $\overline{q}$ endpoints, whenever they are massive. Note that, after the accurate endpoint location has been determined, the corresponding offset correction has to be included as stated in eq. (4.46). A massive system might be formed by one massless q and one massive $\overline{q}$, in which case, eq. (4.48) has to be applied only to the massive $\overline{q}$, while the correction given by the space–time offset in eq. (4.46) has to be used both for q and $\overline{q}$.

A final feature has to be included in the massive space–time implementation, to account for the reduced oscillation period of the endpoint hadrons consisting of massive quarks, which is shorter than for hadrons formed by massless quarks. This discrepancy only affects the estimation of the "late", v_l^h, and "early", v_e^h, definitions of hadron production points, whose expressions, previously given by eq. (4.33) and eq.(4.34), become

$$v_{l/e}^h = \frac{v_i + v_{i+1}}{2} \pm \alpha_{\text{red}} \frac{p_h}{2\kappa} \tag{4.49}$$

where α_{red} is the parameter that accounts for the reduced oscillation period. This parameter is determined in the hadron rest frame by the three-momenta of the quarks forming

the hadron, i.e. $\alpha_{\rm red} = p_0/m_h = |\overrightarrow{p_q}|/m_h = -|\overrightarrow{p_{\overline{q}}}|/m_h$, whose general expression, p_0, is displayed in eq. (4.41). Thus, the reduced oscillation parameter is

$$\alpha_{\rm red} = \frac{\sqrt{\left(m_h^2 - m_q^2 - m_{\overline{q}}^2\right)^2 - 4m_q^2 m_{\overline{q}}^2}}{m_h^2} \tag{4.50}$$

where m_h is the mass of the hadron and m_q and $m_{\overline{q}}$, the masses of the constituent q and $\overline{q}$, respectively.

4.6 Other implementation details

Up until now, only $q\overline{q}$ systems, multiparton systems and gluon-loop topologies have been considered. Although those systems are the most common ones, systems composed of three different quarks or antiquarks are also possible, specially in pp collisions, where two of the hadrons created during hadronization have to be baryons due to the conservation of baryon number. Those systems, known as junction systems, also fragment in the hadronization process and need a special mention in the space–time implementation.

A junction system consists of three different "legs" or subsystems, whose ends correspond to the quarks forming the junction system, which define the flavour content of the system, and the junction proper, where the colour-string legs meet. The hadronization process is most conveniently performed in the rest frame of the junction, in which the total energy of each leg is determined and the two legs with the lowest energy are fragmented first. The fragmentation process in each leg is performed always stepping from the $q/\overline{q}$ end towards the centre of the junction system. The process stops whenever the next step to be taken would require more energy than the energy left in the leg.

Once the two initial legs have fragmented, a q or $\overline{q}$ from the last break is left in each of them. These two partons combined to create a diquark. Together with the third leg and its orginal endpoint $q/\overline{q}$, this diquark defines a final string system, which fragments following the same process described in section 3.4.

In Pythia, the assignment of the space–time location to the hadrons is performed once all the hadrons of the system have been created in the energy–momentum picture. The fragmentation process of a junction system follows an order, which has to be taken into account when assigning the space–time locations to the breakups, already generated. This is done by considering the three different legs to be three different systems and dealing with them in the same order as they fragmented, starting from the leg with the lowest energy. Apart from this, the calculation of the space–time locations and all the implementations presented so far are performed in the same way as in a $q\overline{q}$ system.

Low-invariant-mass systems also hadronize, following the same fragmentation process as in systems with higher invariant masses. The only difference takes place when the invariant mass of the system is so low that it can only form one hadron. In such cases, the

"early" description of hadron production point corresponds to the origin of the $q\overline{q}$ system, i.e. $v_i^h = (\ 0;\ 0,\ 0,\ 0)$. Note that the smearing in transverse space will give rise to negative squared invariant times in such cases, which is not considered to be a problem due to the pancake picture (see section 2). The "middle" and "late" definitions are calculated from the four-momentum of the hadron p_h as $v_i^h = p_h/2$ and $v_i^h = p_h$, respectively.

5 Hadron density studies

In this section, the space–time implementation is used to study the space–time hadronic density in different situations. First, we show that the three different definitions of hadron production points (see section 4.1) can be regarded as a measure of uncertainty. This is done by analysing the longitudinal and transverse spectra of hadron production in section 5.1. Then, the temporal evolution of hadron production is studied in section 5.2, both for single string $q\overline{q}$ systems and for pp collisions, for which the radial evolution is also analysed.

Once the general evolution of hadron production is understood, the close-packing feature is addressed. First the hadronic density in the central region of the collision is tackled in section 5.3, showing the presence of overlapping hadrons. To extend the study to all the regions and determine the effects of close-packing for different multiplicities, a similar analysis is performed as a function of hadronic multiplicity for pp collisions at $\sqrt{s} = 13$ TeV, as presented in section 5.4. This issue is examined further in section 5.5, where the number of hadrons overlapping with a newly generated one are analysed for different multiplicity ranges, as well as the relation between close-packing and the hadron transverse momentum.

5.1 Longitudinal and transverse distributions

In section 4.1, the three definitions of hadron production point were presented. In order to prove that the different results drawn from the definitions can be used as a measure of uncertainty, the longitudinal and transverse space–time distributions are studied. The transverse spectrum is represented using the variable $r = \sqrt{x^2 + y^2}$, while for the longitudinal spectrum a "space–time" rapidity, y_τ, is used

$$y = \frac{1}{2}\log\left(\frac{E + p_z}{E - p_z}\right) \longrightarrow y_\tau = \frac{1}{2}\log\left(\frac{t + z}{t - z}\right) \tag{5.51}$$

where t, x, y and z are the coordinates of the hadron production points. Note that the longitudinal variable is dimensionless while the transverse one is expressed in units of fermi (fm). Although formally unrelated, the dynamics of string fragmentation introduces a strong correlation between y and y_τ, as illustrated in Figure 16. The spread from the diagonal comes from a number of effects, such as the probabilistic fragmentation process, given by eq. (3.15), and hadronic decays.

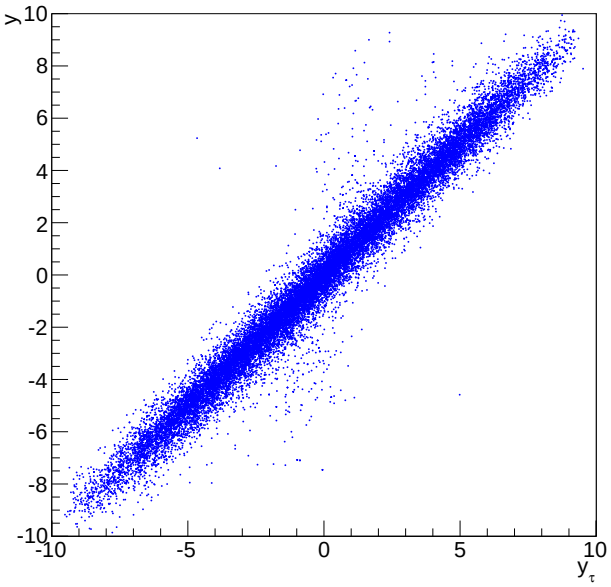


Figure 16: Correlation between rapidity, y, and the equivalent space–time rapidity, y_τ, for all hadrons in 100 pp collisions at $\sqrt{s} = 13$ TeV.

Figures 17 and 18 display the longitudinal and transverse spectra for pp collisions at $\sqrt{s} = 13$ TeV given by the "early", "middle" and "late" definitions of hadron production points, represented in green, red and blue, respectively. In the same figures, the spectra for a single string, whose CM energy is $\sqrt{s} = 13$ TeV, using the "middle" definition are also illustrated in black. In both cases primary and secondary hadrons are taken into account and all the distributions have been normalized to the number of events simulated and the bin width.

As can be seen in Figure 17, the longitudinal spectra given by the different definitions are extremely similar. The largest disagreement among definitions can be seen for values around $y_\tau \approx 0$, where the spectra of the "early" definition peaks at a higher value than in the other two cases. Another small discrepancy is observed for larger y_τ values, where especially the "early" description seems to behave slightly different than the other two definitions. The reason for those disparities is the fact that the "early" alternative maximizes the extreme behaviour of hadron production. Apart from that, the results obtained from the three alternatives are similar enough to consider the differences among them as an uncertainty measure.

Similar conclusions are drawn from the transverse spectra, shown in Figure 18. The Gaussian shape, especially in the $q\overline{q}$ spectrum, is a consequence of the smearing in transverse space, which follows a Gaussian distribution, as explained in section 4.4. This Gaussian shape is distorted in the pp transverse spectra due to the large number of strings that constitute pp events, each of them stretching in a different way between the partons

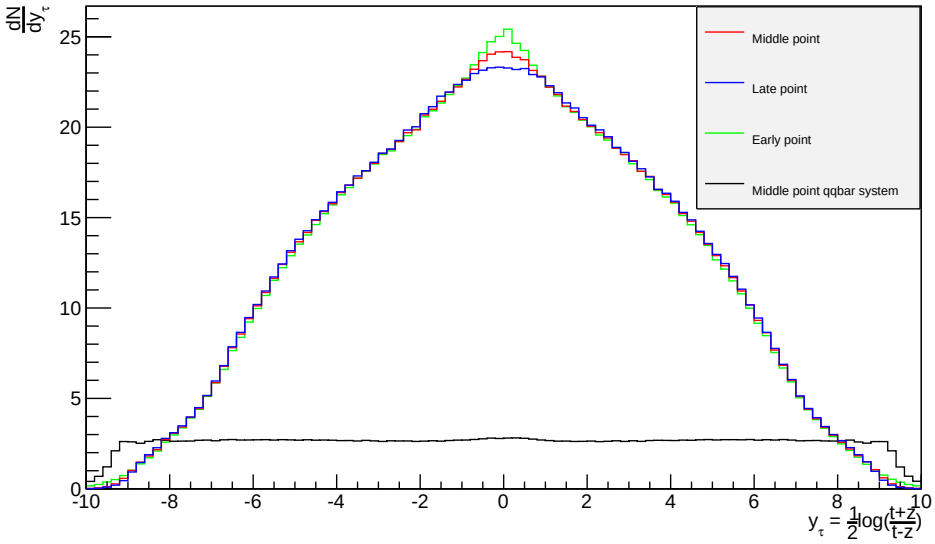


Figure 17: Longitudinal spectra for pp collisions and $q\bar{q}$ systems, both at $\sqrt{s} = 13$ TeV.

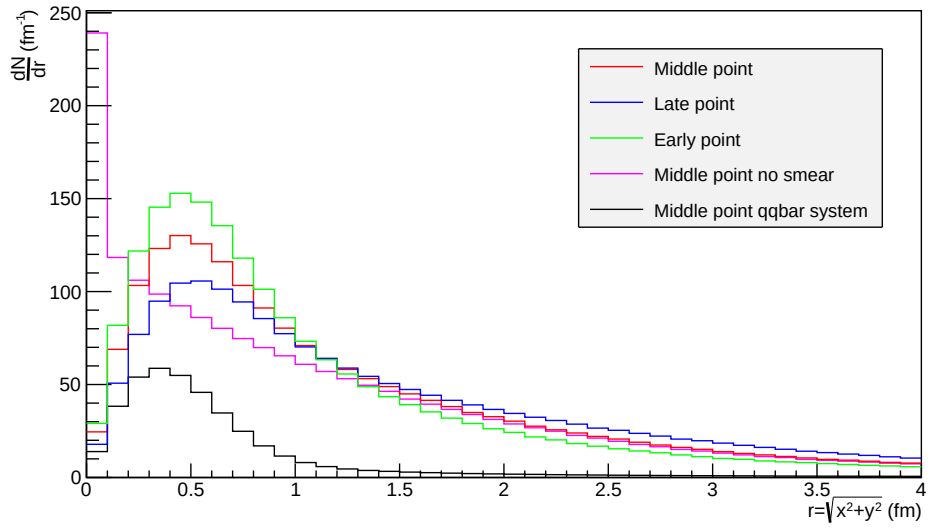


Figure 18: Transverse spectra for pp collisions and $q\bar{q}$ systems, both at $\sqrt{s} = 13$ TeV.

from hard collisions, parton showers and beam remnants. For the same reason, the $q\bar{q}$ transverse spectrum drops faster and peaks closer to zero than the pp spectra. As in the longitudinal case, the pp transverse spectrum given by the "early" definition shows a more extreme behaviour than the other two definitions. Although the discrepancy among the results for the three definitions is greater than in the longitudinal spectrum, we conclude that the three definitions for hadron production points are equivalent and can be use as an uncertainty measure.

To illustrate the effect of the transverse smearing, the transverse spectrum for 13 TeV pp collisions given by the "middle" definition without including smearing is also displayed in Figure 18. In that case, the transverse spectrum peaks at $r = 0$ fm and falls for increasing radii, indicating the presence of some strings aligned with the beam axis, while other stretch in diverse ways along the transverse direction. Although the spectrum for small radii changes when including transverse smearing, displacing the peak at $r = 0$ fm to a larger value of $r \approx 0.5$ fm, the behaviour for large radii is unchanged thanks to the assumption that the transverse location of the breakups follows a Gaussian.

5.2 Temporal and radial evolution of hadron production

Once the transverse and longitudinal spectra have been displayed, let us focus in this section on the temporal and radial evolution of hadron production in pp and $q\overline{q}$ systems. Note that the space–time quantities are expressed in units of fermi (fm), if not stated otherwise.

The number of hadrons in the system is displayed as a function of time for a single string system at $\sqrt{s} = 20$ GeV in Figure 19, and for pp collisions at $\sqrt{s} = 13$ TeV in Figures 21 and 22. The red curve corresponds to the number of primary hadrons, or hadrons formed during the hadronization process, while the green curve represents the number of secondary hadrons in the system, generated from particle decays. The total number of hadrons in the system, illustrated in blue, is given by the sum of the primary and secondary hadrons present in the system. The brown curve represents the number of final hadrons in the system, or hadron regarded as stable in event generators. Finally, the black curve depicts the number of hadrons located at the longitudinal centre of the collision, i.e. at a location $|z| < 0.5$ fm, a behaviour that will be discussed later in section 5.3.

As illustrated in Figure 19, the number of primary hadrons peaks at $t = 10$ fm for a 20 GeV simple $q\overline{q}$ system formed by massless quarks in the CM frame. This effect arises from the fact that the system is completely hadronized at $t > 10$ fm and, thus, no more primary hadrons can be formed. The maximal string length for this system is $l = 10$ fm and, because massless quarks travel at the speed of light, it is reached at $t = 10$ fm, when the hadronization process comes to an end. The number of secondary hadrons also peaks at $t = 10$ fm due to the decays of short-lived primary hadrons generated in the hadronization process. Since some of the primary hadrons decay into other particles, including hadrons, the number of secondary hadrons increases as the number of primary hadrons decreases for $t > 10$ fm. Another important effect observed is the fact that the hadronic outflux from the central region is larger than the influx, showing that hadrons tend to move further out in the longitudinal direction. This is the reason why the number of hadrons located in the central region is only relevant for times $t \leq 4$ fm. This condition is required in the analysis performed in the next section.

Note that there are no hadrons in the system up until $t \approx 0.5$ fm, because it takes some time for hadronization to start, since the string has to begin stretching out before it can

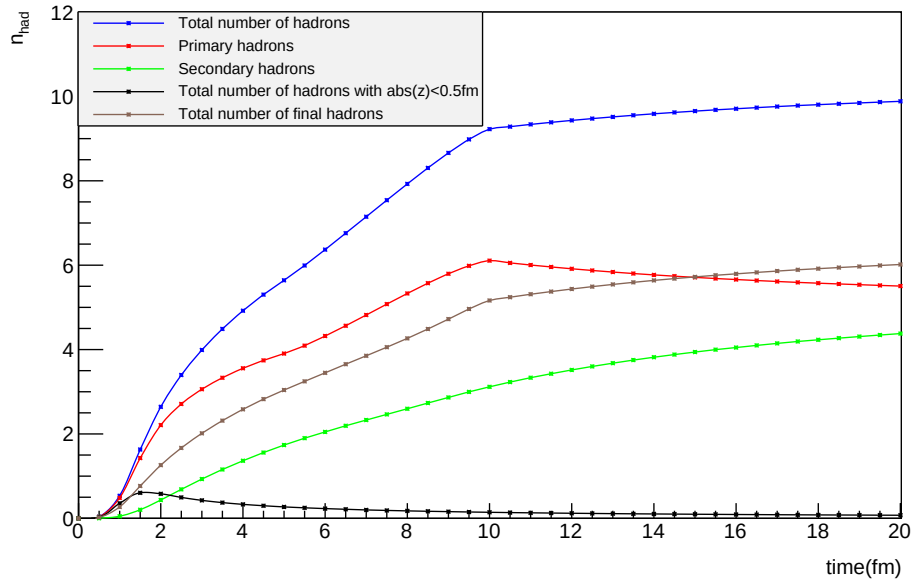


Figure 19: Hadron number per event as a function of time for a simple $q\overline{q}$ system formed by massless quarks in the CM frame with $\sqrt{s} = 20$ GeV.

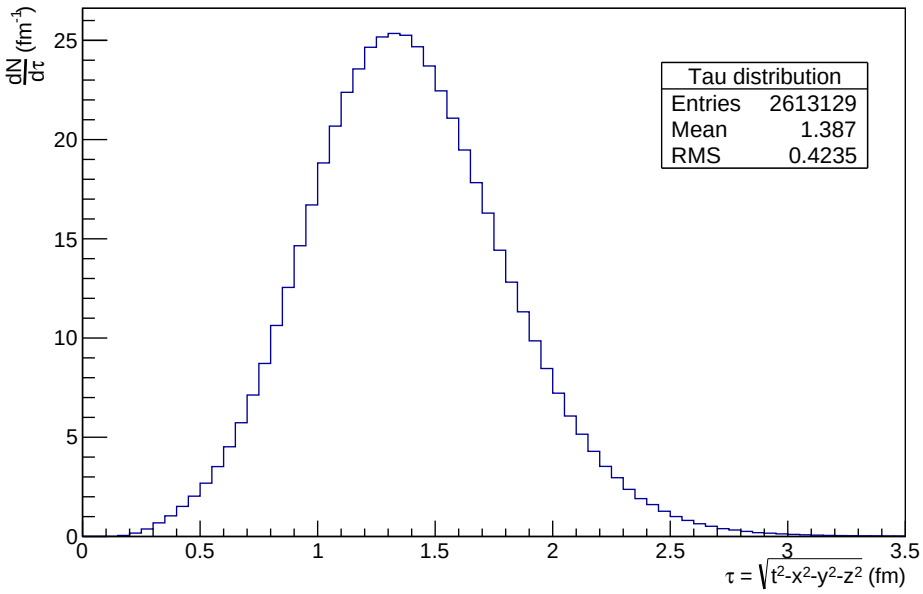


Figure 20: Invariant time τ distribution of primary hadrons generated in $q\overline{q}$ systems.

begin fragmenting. This feature is illustrated in Figure 20, which displays the invariant time distribution of the primary hadrons in the $q\overline{q}$ system presented earlier in this section, normalized to the number of events and the bin width. By default, the parameters a and

b in eq. (3.15) and eq. (3.17) are set to $a = 0.68$ and $b = 0.98$ GeV^{-2} [23], giving rise to a suppression of the small Γ values and an expectation value of $\langle\Gamma\rangle = (1+a)/b \approx 1.7$ GeV^2, determined from eq. (3.17). Due to the relation between Γ and the invariant time τ, given by eq. (3.12), the τ expectation value is $\langle\tau\rangle = \sqrt{\langle\Gamma\rangle}/\kappa \approx 1.3$ GeV, as also illustrated in Figure 20. The relation between Γ and τ also implies the suppression of small τ values, as shown in Figure 20, since $P(\Gamma) \approx \Gamma^a d\Gamma \propto \tau^{2a}\tau d\tau = \tau^{2a+1}d\tau$ for $\tau \to 0$. Because those aspects are typical of the hadronization process, a similar behaviour is also observed in pp collisions.

The time evolution of hadron production in 13 TeV pp collisions is displayed in Figure 21 for times $t \leq 20$ fm. Although the qualitative behaviour is similar to the hadron production in $q\overline{q}$ systems, the temporal evolution is smoother and the number of hadrons generated per unit time increases more rapidly in pp system than in the simple $q\overline{q}$ system. These effects are direct consequences of the presence of several string systems in pp events, whose strings are more spread out, possibly extending out to 6500 fm.

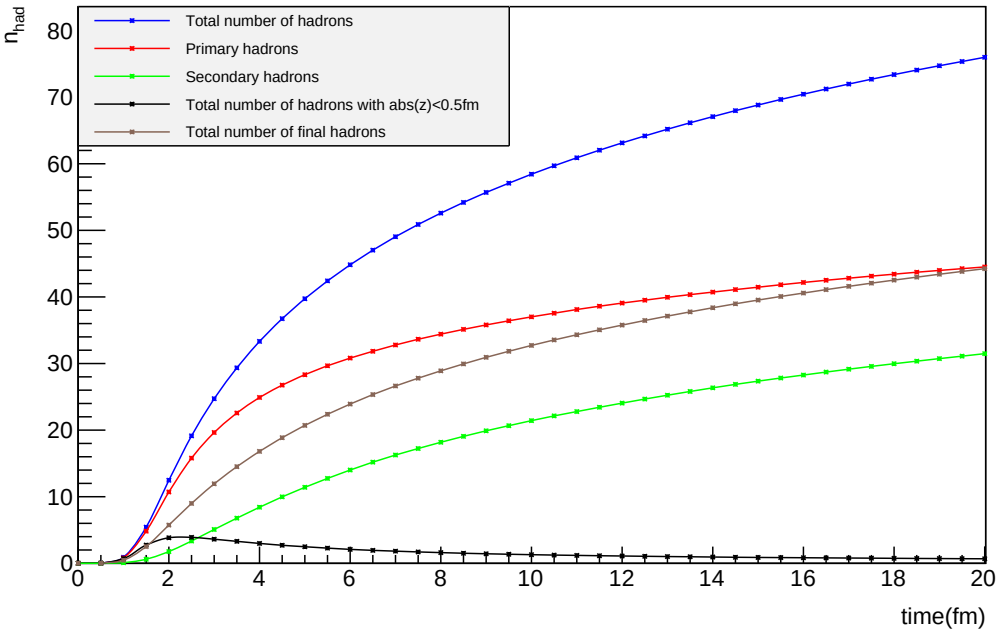


Figure 21: Hadron number per event as a function of time, up until $t = 20$ fm, for pp collisions at $\sqrt{s} = 13$ TeV.

To understand the whole time evolution of hadron production in 13 TeV pp collisions, Figure 22 shows it in a logarithmic scale extending from 1 fm to 10^{15} fm, equivalent to a distance of 1 meter. As in the case of the $q\overline{q}$ system, the total number of primary hadrons increases until hadronization is over ($t \approx 10^3$ fm), which occurs at later times than in the $q\overline{q}$ systems due to the presence of several string systems, some of which are much more

energetic than the $q\overline{q}$ presented before. As time passes, unstable primary hadrons decay into secondary hadrons, generating an increase of the number of secondary hadrons in the system, while the number of primary hadrons decreases. The significant drop in the total number of hadrons at $t \approx 10^8$ fm, also visible in the number of primary and secondary hadrons, represents the electromagnetic decay of the π^0 hadrons present in the system into photons, $\pi^0 \rightarrow \gamma + \gamma$. Although the lifetimes of hadrons composed of s, b and c quarks in their rest frame are of order of some millimeter or centimeter, their decays are still ongoing at times $t = 10^{15}$ fm, since the hadrons are in a boosted frame with respect to their rest frame. All the unstable primary and secondary hadrons will decay until the only hadrons left in the system are stable hadrons, meaning that the blue and yellow lines will meet. As mentioned in section 2, the decays of $\mu^{\pm}$, $\pi^{\pm}$, $K^{\pm}$, K^0_L, n and $\overline{n}$ are not considered or else only a few p and $\overline{p}$ would remain per event, since p and $\overline{p}$ are the only truly stable hadrons.

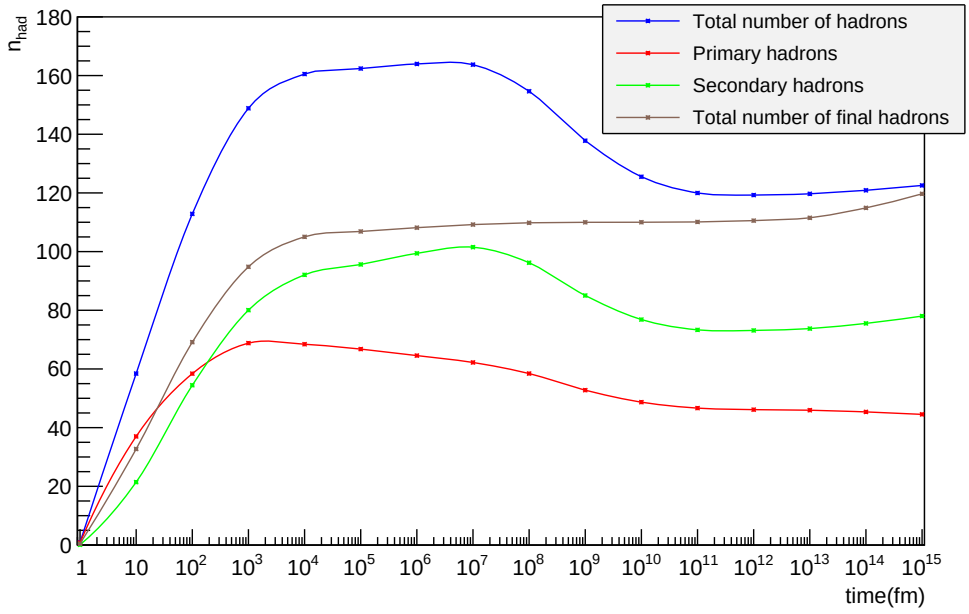


Figure 22: Hadron number per event as a function of time for 13 TeV pp collisions.

Alternatively, the radial evolution of hadron production for 13 TeV pp collisions is displayed, in a logarithmic scale extending from $r = 1$ fm and $r = 10^{15}$ fm $= 1$ m, in Figure 23. Overall, the diverse effects visible in the time evolution of the system are also present in the radial evolution. Some of the primary hadrons are generated close to the collision centre, where $r < 1$ fm, meaning that the distribution of primary hadrons does not fall to zero for $r < 1$ fm. The same feature is observed for secondary hadrons, since some of them are created close to the collision centre, generated in the decays of short-lived primary hadrons, such as ρ. As in the time evolution, the significant drop at $r \approx 10^6$ fm in the primary, secondary and total number of hadrons is due to the electromagnetic decay $\pi^0 \rightarrow \gamma + \gamma$. The number of final stable hadrons increases slightly at $r \approx 10^{13}$ fm, implying

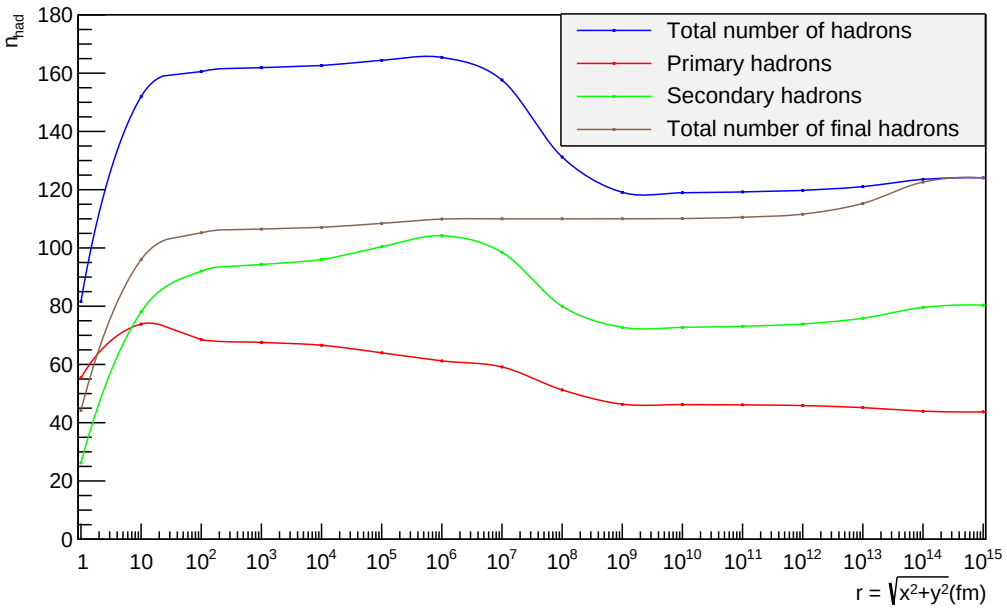


Figure 23: Hadron number per event as a function or radius for 13 TeV pp collisions

that almost all weak decays of the hadrons composed of s, c and b quarks take place before $r \approx 10^{15}$ fm, making the number of stable final hadrons and the total number of hadron in the system coincide. Although this aspect could have been inferred form the time evolution in Figure 22, the total number of hadrons and the stable hadrons do not coincide at time $t \approx 1$ m/c since hadrons can have a very large longitudinal motion.

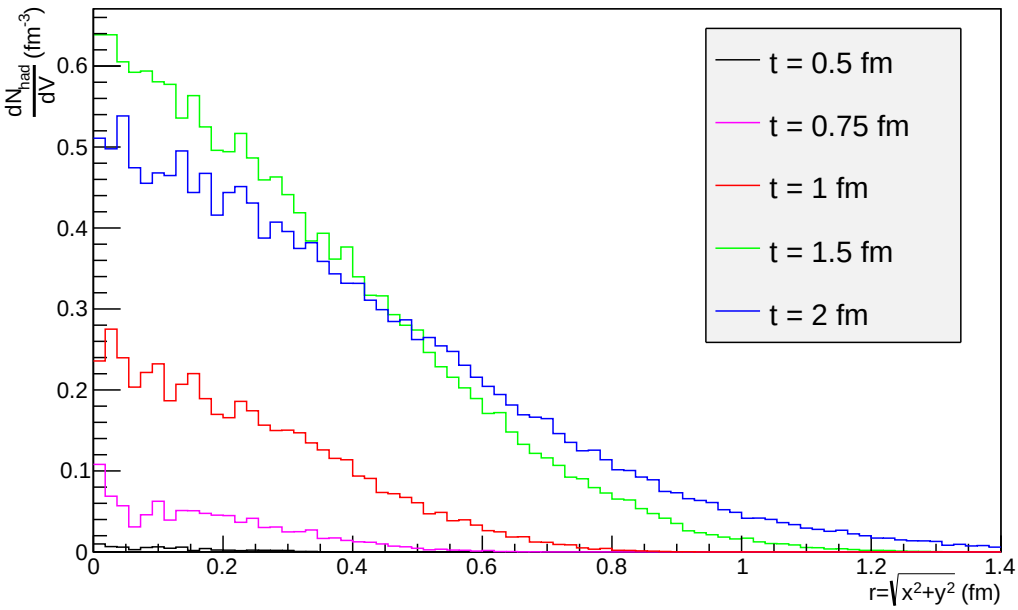
5.3 Close-packing of hadron production in the central region

The space–time density of hadron production, dN/dV, can be analysed in terms of different measures, which must be Lorentz invariant to give rise to frame independent conclusions. Since Lorentz invariant measures tend to hide the time aspect, to begin with the analysis of hadronic density, the hadron density is studied in this section in a specific region. Particularly, the time evolution of the hadronic density is studied in the longitudinal central region, defined such that $|z| \leq 0.5$ fm, a choice that is approximately Lorentz invariant under longitudinal boosts in the central rapidity plateau. For that region, the hadronic density can be expressed in cylindrical coordinates as

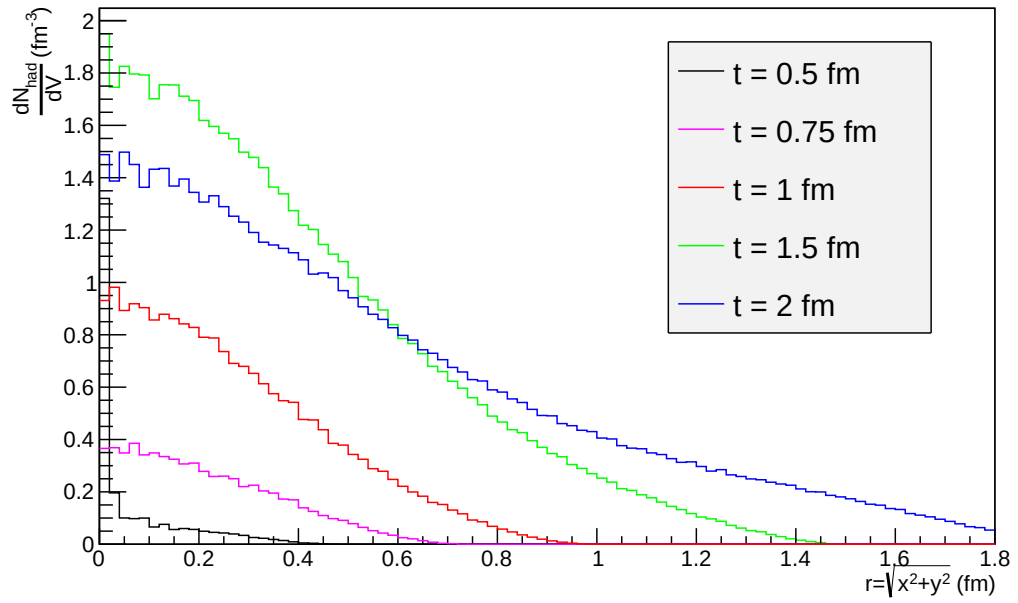
$$\left.\frac{dN}{dV}\right|_{|z|\leq 0.5} = \left.\frac{dN}{dxdydz}\right|_{|z|\leq 0.5} = \frac{dN}{dxdy} = \frac{dN}{2\pi r dr} \tag{5.52}$$

giving a measure of the hadronic densities as a function of radius.

Figures 24a and 24b show the distribution of the hadronic densities in the central region as a function of the radius, determined by eq. (5.52), and for different times, for $q\bar{q}$ systems



(a) 20 GeV single string systems



(b) 13 TeV pp collisions

Figure 24: Hadronic density as a function of the radius for different constant times, for a central slice $|z| < 0.5$ fm.

at $\sqrt{s} = 20$ GeV and 13 TeV pp collisions, respectively. Only times up to $t = 2$ fm have been considered, since the total number of hadrons in the central region peaks at

$t \approx 2$ fm both in $q\overline{q}$ and pp systems, as illustrated in the previous section, particularly in Figures 19 and 21. In both distributions, primary and secondary hadrons have been taken into account. The hadron density at times $t = 0.5$ fm is extremely low in both systems, since they hardly have had time to start hadronizing yet. From this point on, hadrons are generated from hadronization and particle decays, increasing the value of the hadron density in the central region. The maximal value of the hadronic density in the region closer to the origin is obtained at $t = 1.5$ fm, both in pp collisions and $q\overline{q}$ systems, a value that relates to typical hadronization time scales. From this time on, the hadron density spreads in the longitudinal and transverse directions, illustrating the motion of the hadrons, which are created at the collision centre and moving away from it.

Another important feature that can be deduced from Figure 24b is the close-packing of hadrons near to the centre of the collision. To demonstrate this, consider the typical hadron volume to be defined by the proton volume, whose radius is $r_p = 0.87$ fm [22], thus $V_h = (4\pi r_p^3)/3 \approx 2.76$ fm^3. The maximum value of the hadronic density for pp collisions, reached at time $t = 1.5$ fm for radii $r \approx 0$ fm, is $dN_{\text{had}}/dV = 1.8$ fm^{-3}, implying that five hadrons overlap in the centre of the collisions, disregarding Lorentz contraction. Therefore, the generation of a new hadron from the hadronization process is affected by the presence of other hadrons overlapping with it. In the following sections the close-packing of hadrons is analysed as a function of event multiplicity.

5.4 Hadron production at different multiplicities

In the previous analysis, we focused on the longitudinal central region, leading to the first evidence of the presence of close-packing in hadron production. Although the hadronic density definition is reasonable for a central slice with limited longitudinal motion, the analysis cannot be extended to all the regions because the results will be frame dependent. In order to perform a Lorentz invariant analysis of the space–time hadron production including all regions, the Lorentz invariant measure d^3x/t is used in this section. Since the number of particles in an event, N, is frame independent, a Lorentz invariant measure of the hadron density can be defined as

$$t\frac{dN}{d^3x} = \frac{dN}{d^2r\frac{dz}{t}} = \frac{dN}{\pi dr^2 dy_\tau} \rightarrow \frac{N}{\pi r_m^2 \Delta y_\tau} \tag{5.53}$$

where Δy_τ is the space–time rapidity range of the event and r is the radius at which particles in that event are created. As we want to study the typical average behaviour of hadron production, we consider r_m to be the median radius of the set of hadrons created in the event, to avoid a strong dependence on the extreme values. For similar reasons, Δy_τ is determined as the extent of the space–time rapidity between two y_τ values at which dN/dy_τ is half of its maximum value, given by dN/dy_τ at $y_\tau = 0$. Note that cylindrical coordinates have been used, i.e. $dxdy = rdrd\varphi = \pi dr^2$. Since we are interested only in the hadron production, we consider N to be the number of hadrons in the event, a variable that can be understood as the event hadronic multiplicity.

Note that the concept of hadronic multiplicity applied to these studies is different than its experimental definition. In experimental particle physics, the hadronic multiplicity of an event is defined as the number of charge hadrons in that event, since they can be associated with tracks in the detector components, while neutral hadrons are detected as energy deposits. For the study of hadron production, all hadrons are equally important, no matter if they are charged or not. Consequently, the event hadronic multiplicity in our studies is defined as the total number of hadrons in an event, subjected to some constraints.

In this analysis, the event hadronic multiplicity has been determined taking into account primary and secondary hadron and including some specific constraints. Since we are interested in the hadronization process, only strong decays should be taken into account in our analysis. For that reason, the hadrons whose lifetimes are $\tau \geq \tau_{\pi^0} \approx 2.55313 \times 10^{-5}$ mm/c [22] are regarded as stable, since their decays are electromagnetic or weak. Similarly, decay products generated at radii $r > 10$ fm are not taken into account in this analysis. In order to include the space–time location of the hadrons and their decay products without accounting for more hadrons than the final hadrons of the event, all secondary hadronic decay vertices enter with a weight one less than the hadronic multiplicity of the decay, thus compensating for the already bookkept decaying hadron.

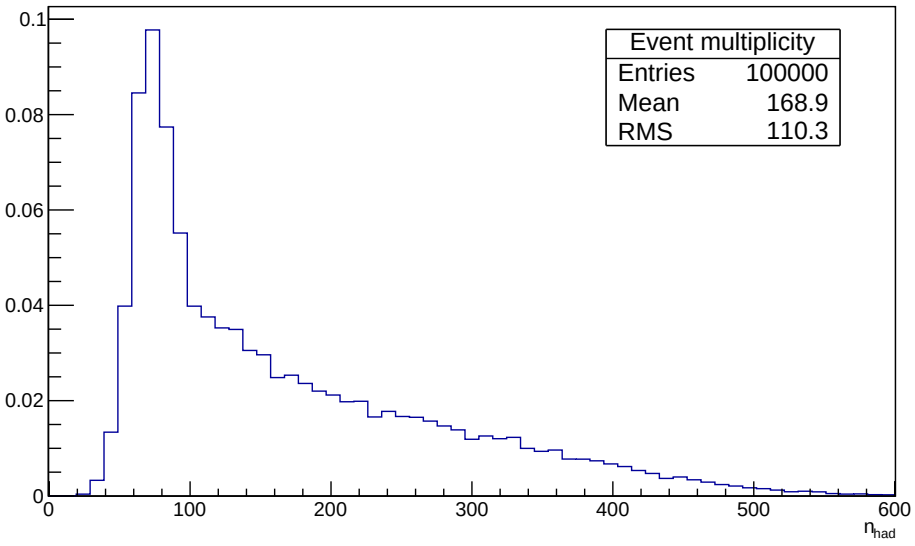


Figure 25: Hadron multiplicity distribution for 13 TeV *pp* collisions.

The event hadronic multiplicity distribution for 13 TeV *pp* collisions simulated in Pythia is displayed in Figure 25, subjected to all the constraints presented and normalized to the number of total events simulated. As can be seen, most of the events generated in Pythia have multiplicities between $n_{\text{had}} \approx 40$ and $n_{\text{had}} \approx 120$, while the average multiplicity value is at $n_{\text{had}} = 169$. Events with multiplicities $n_{\text{had}} > 400$ are generated at a much

lower rate. For our purposes, ten multiplicity ranges are defined, based on the multiplicity distribution, by simulating one hundred thousand 13 TeV *pp* collisions and requiring each multiplicity range to contain around ten thousand events.

Before moving forward with the analysis of hadronic density as a function of hadron multiplicity, the space–time longitudinal and transverse spectra of hadron production in 13 TeV *pp* collisions are presented for different multiplicity ranges in Figures 26 and 27, respectively, normalized to the number of events in each range and the bin width. For the sake of clarity, some intermediate bins are left out of the figures. In Figure 26, the space–time rapidity range decreases with multiplicity because more hadrons are produced in the central region at higher multiplicities. The hadron production at the endpoints is suppressed for increasing multiplicity due to energy-momentum conservation, since the energy to create more central hadrons has to be taken from the beam remnants.

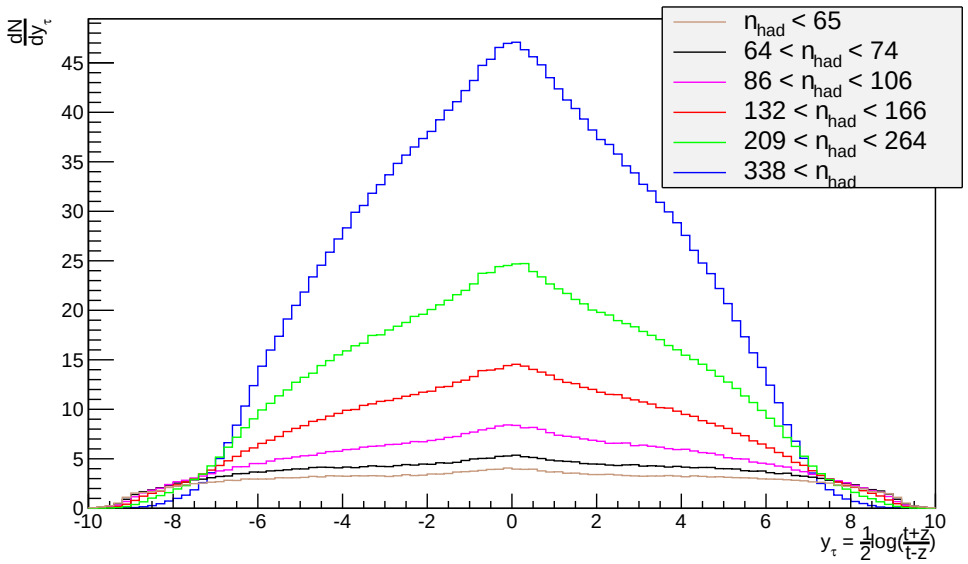


Figure 26: Longitudinal spectra for 13 TeV *pp* collisions and different multiplicity ranges.

Similar conclusion are drawn for the transverse spectra shown in Figure 27. Since at high multiplicities more hadrons are generated, the number of hadrons per unit radius increases with multiplicity. The median and average radii also increase with multiplicity, which could be an effect of colour reconnection, a property by which partons from disconnected subprocesses are tied together, generating fewer string pieces and more MPIs than otherwise.

The effect of colour reconnection is analysed in Figure 28, which displays the median radii as a function of the median hadronic multiplicity, n_{had}, of each multiplicity range. The red and blue curves represent the values obtained with and without including colour

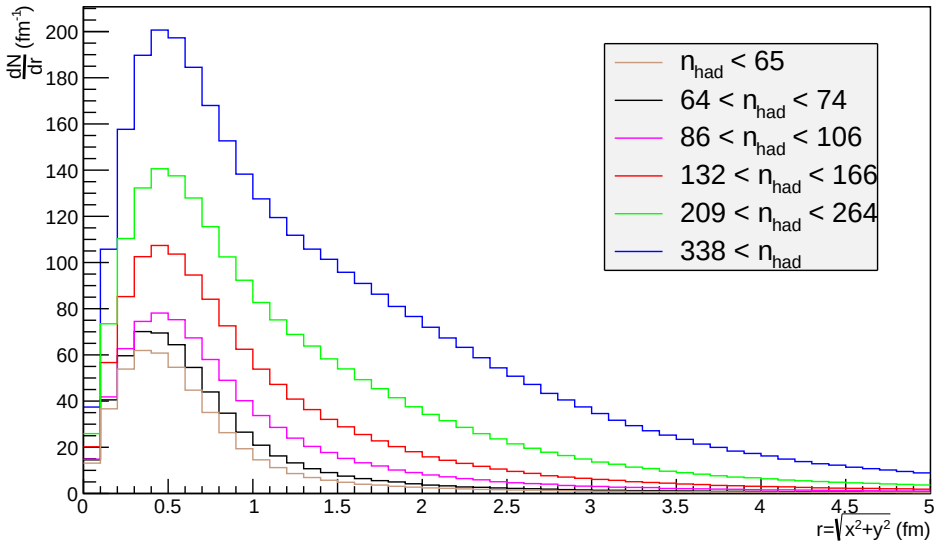


Figure 27: Transverse spectra for 13 TeV *pp* collisions and different multiplicity ranges.

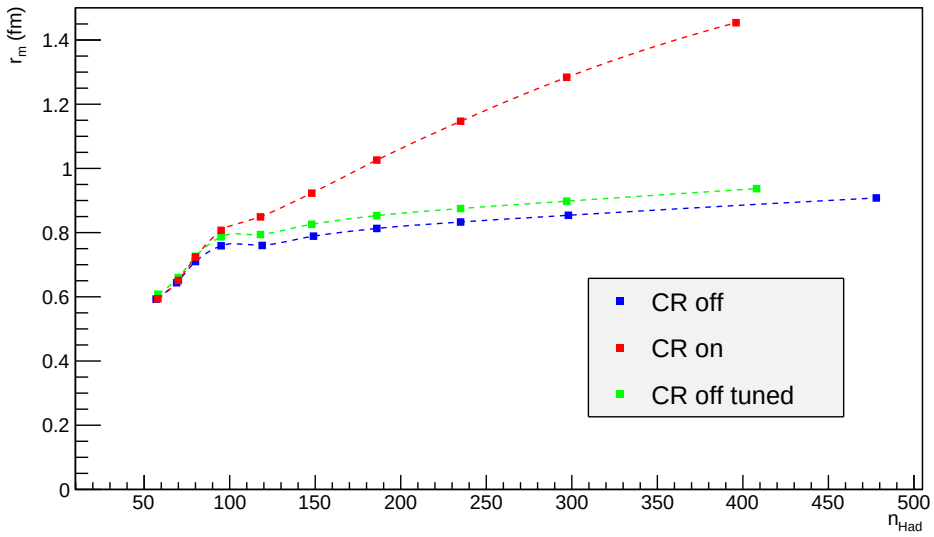


Figure 28: Median radii as a function of multiplicity for 13 TeV *pp* collisions. The red curve corresponds to the approach with colour reconnection, while the blue and green curves represent the model without colour reconnection and a tuned model without colour reconnection, respectively.

reconnection, respectively. Although the multiplicity ranges in both approaches are equally defined, the median n_{had} value for the last range differs for both configurations, due to the increase in multiplicity when colour reconnection is not included. Consequently, a new

set of values is included in Figure 28 in green, corresponding to a CR off model tuned to have the same median multiplicity as when colour reconnection is on. This adjustment is performed by reducing the number of low-p_T MPIs in the CR off scenario. Although both CR off models follow the same trend, the values at high multiplicity scales are slightly different due to the fact that the hadrons in the tuned model have a slightly higher p_T, as the p_T of the average MPI is higher, and particles decay further out, meaning that the production vertices take place at larger space–time scales. The discrepancy between the values with and without colour reconnection grows as multiplicity increases since more and larger strings are present at higher multiplicities when colour reconnection is not included, which are more aligned with the longitudinal direction.

To determine the hadron production behaviour as a function of multiplicity, the values n_{had}, r_m and Δy_τ have to be determined in each multiplicity range by taking into account all the events inside each range. As said before, r_m is the median radius of the set of hadrons inside that range, where radius is defined as $r = \sqrt{x^2 + y^2}$. Similarly, Δy_τ is determined in each multiplicity range from the set of hadrons inside the range as described previously in this section. From that approach, the endpoint behaviour is excluded as wanted, since our analysis focuses on the typical hadron production given by more central regions. Finally, n_{had} is defined in each multiplicity range as the median hadronic multiplicity value of all the events in that range.

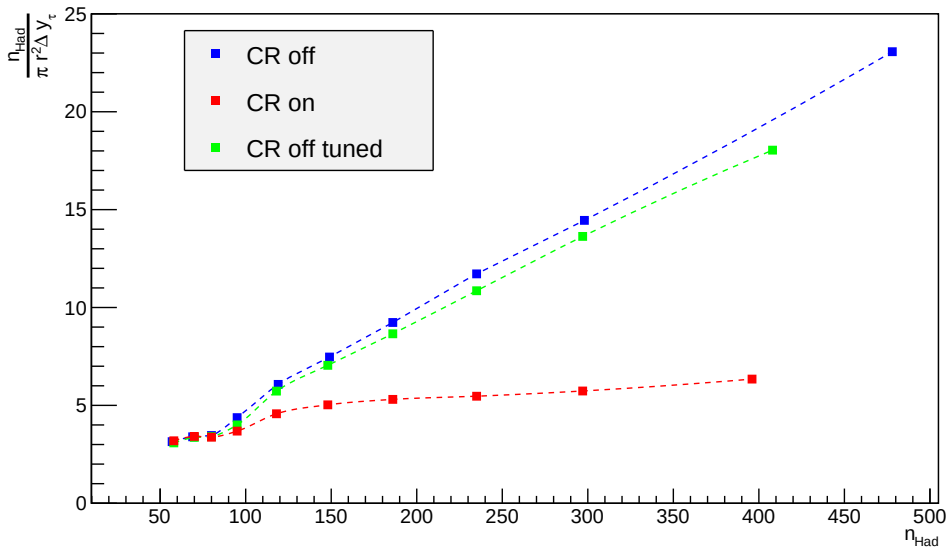


Figure 29: Hadron density as a function of multiplicity for pp collisions at 13 TeV. The red, blue and green curves represent the three different models with and without colour reconnection, also included in Figure 28.

Figure 29 displays the hadron density, defined in eq. (5.53), as a function of hadron multiplicity for 13 TeV pp collisions for the three different scenarios presented in Figure

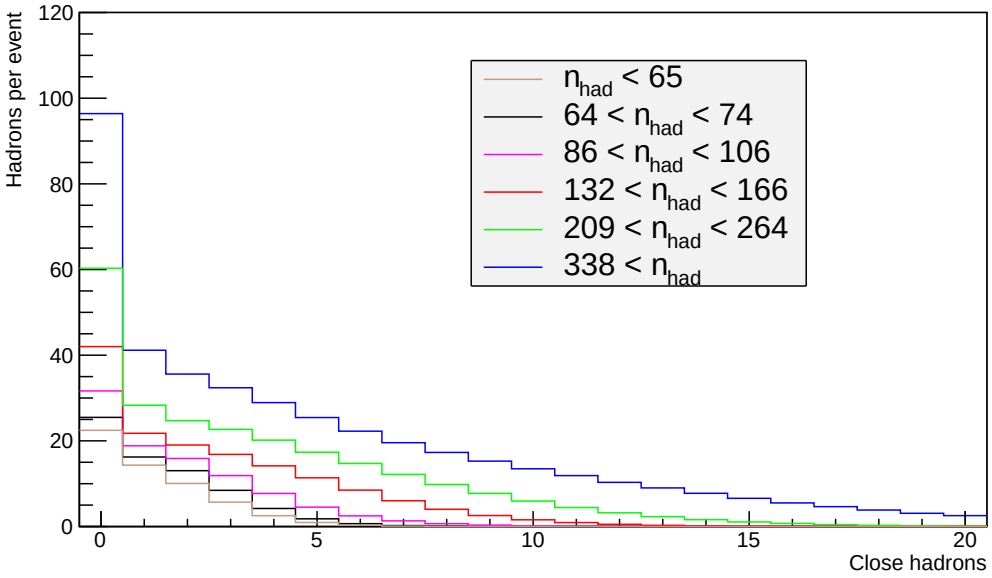
28. Overall, the space–time hadron density increases with hadronic multiplicity. The densities for the two models without colour reconnection are significantly higher at large multiplicities than the approach including colour reconnection, an effect arising from the divergent trend in the median radii for the two CR off models, illustrated in Figure 28. In the three cases, the rise in the density is less dramatic for lower multiplicity values while the hadron density increases almost linearly for $n_{\text{had}} > 150$. Even though close-packing is enhanced in high-multiplicity *pp* collisions in the three models, its effect is more significant when colour reconnection is not taking into account. Yet, even with the default CR on scenario, it is clear that particles are produced closely packed, i.e. with a significant overlap.

5.5 Close-packing analysis in hadron rest frame

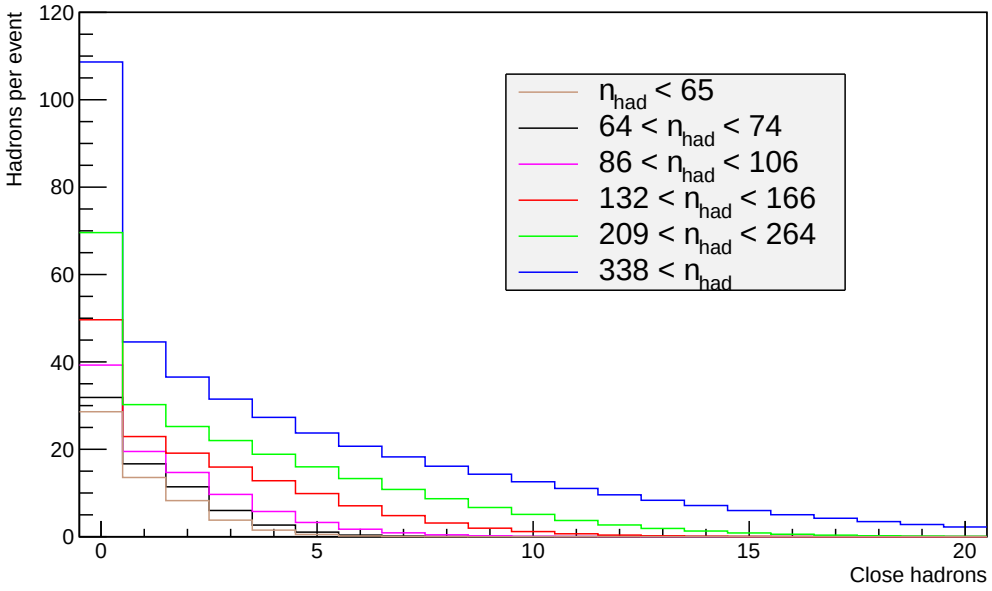
The close-packing of hadron production is further analysed in this section by focusing on the hadronic overlap. This analysis is performed by regarding hadrons as three-dimensional spheres whose radius is the proton radius, $r_p = 0.87$ fm [22]. In order for this analysis to be Lorentz invariant, the number of hadrons overlapping with a given hadron are counted in the rest frame of the given hadron. To illustrate the process followed, consider a hadron h_1 generated at time t_{h_1}, where t_{h_1} is defined in the rest frame of hadron h_1. The rest of the hadrons in the system are boosted to the rest frame of h_1, where only the hadrons created at times $t \leq t_{h_1}$ and which have not decayed at $t = t_{h_1}$ are taken into account. Since the exact location of each hadron at $t = t_{h_1}$ can be determined from the hadron production point, defined in section 4.1, and its four-momentum, the distances of each hadron with respect to h_1 when the latter is created can be calculated. If such distances are shorter than $2r_p$, the hadrons are considered to overlap, implying that the production of h_1 could be affected by the presence of these other hadrons. Note that Lorentz contraction is not taken into account. In this analysis, two cases are studied, including and excluding the adjacent hadron on each side along the string, since any effect of such overlaps is already taken into account in the tuning of the fragmentation process, in eq. (3.15) and eq. (3.17). This procedure is followed for each of the hadrons generated in the event.

The hadron overlapping distributions in the rest frame of each hadron for *pp* collisions at $\sqrt{s} = 13$ TeV are displayed in Figures 30a and 30b, including and excluding adjacent hadrons and for different hadronic multiplicity ranges, where the definition of hadron multiplicity is presented in section 5.4. As in the previous analysis, each distribution is normalized to the number of events considered in each multiplicity range, and some intermediate bins are left out of the figures for the sake of clarity. As expected, more hadrons overlap when the adjacent hadrons are taken into account. Although close-packing in hadron production also takes place in low-multiplicity *pp* collisions, the number of hadrons overlapping with a newly generated one is not significantly high. That is not the case for high-multiplicity *pp* collisions, where close-packing often arises with a significant number of nearby hadrons, likely leading to unknown effects that are not taken into account in Pythia.

In the analysis so far, hadrons have not been separated by any kinematics variable.



(a) Including adjacent hadrons



(b) Excluding adjacent hadrons

Figure 30: Hadron overlap for different multiplicity ranges for 13 TeV pp collisions.

Generally, particles produced at large transverse momenta are not expected to suffer from close-packing: even if parton showers can generate many partons from each initial high-p_T parton, these daughter partons are spread widely in momentum space. Therefore, the fragmenting strings stretched between them also will have a modest overlap, unlike the

accretion of low-p_T strings from multiple soft MPIs. In order to isolate this feature, we study the hadron overlap as a function of the hadron transverse momentum, by determining the number of nearby hadrons in the hadron rest frame. In this case, the adjacent hadrons are excluded when determining the number of overlapping hadrons.

The average number of nearby hadrons in the hadron rest frame as a function of the hadron p_T is illustrated for 13 TeV pp collisions in Figure 31, for soft and hard QCD in red and blue, respectively. For the hard QCD approach, the cutoff for the minimum jet transverse momentum is set to 100 GeV. Note that, although the number of nearby hadrons have been counted hadron by hadron, this analysis gives an approximate average measure of hadronic overlap in the rest frame. Even though the number of nearby hadrons is higher for hard QCD processes than for soft QCD, since more particles are generated in hard processes, the trend is the same in both cases. The number of nearby hadrons peaks at $p_T \approx 0.5$ GeV in both cases, although at a higher value for hard QCD processes because such processes are biased towards central pp collisions, which are more likely to produced a high p_T jet, but also generate more soft MPIs. To the right of the peak, both distributions drop dramatically until $p_T \approx 4$ GeV, whereafter they continue to drop rather slowly. This trend implies that close-packing only affects low-momenta hadrons, particularly hadrons with transverse momentum $p_T \leq 3$ GeV, meaning that the hadron production properties are different in different regions.

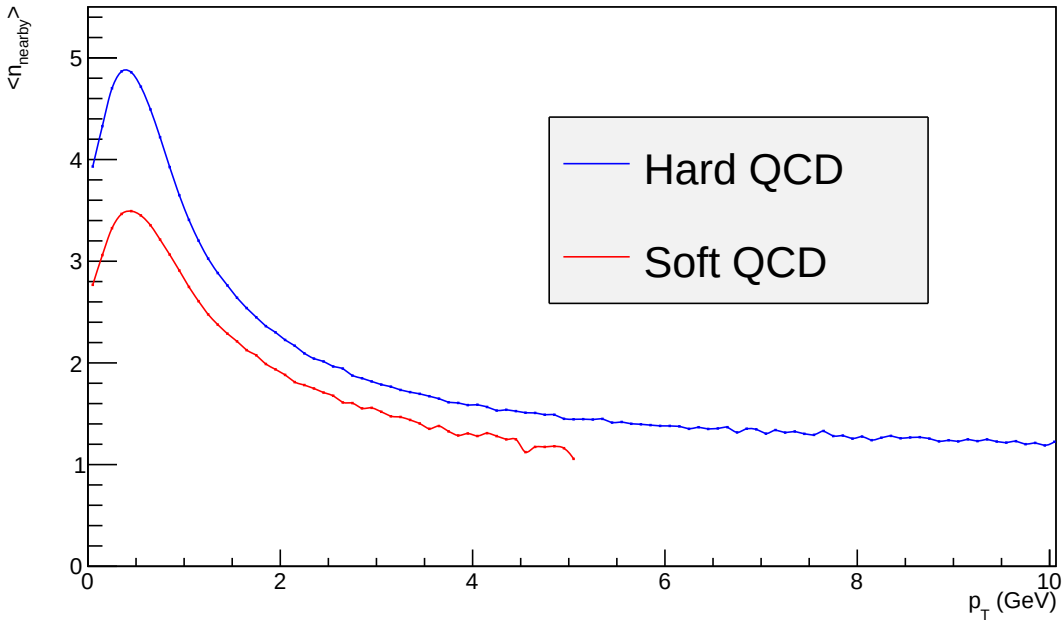


Figure 31: Average number of nearby hadrons inside a jet as a function of the jet transverse momentum for 13 TeV pp collisions. The red and blue distributions illustrated the soft and hard QCD processes.

6 Summary

The aim of this thesis was to start addressing the unexpected collective behaviour of high-multiplicity pp collisions, typical of heavy-ion collisions and the production of Quark-Gluon plasma, by analysing the hadronic density in terms of different measures and for different multiplicities. To perform these studies, the space–time picture was first implemented in Pythia, by determining the space–time location of the string breakups and establishing three alternative definitions for primary hadron production points. Different situations have been tackled, such as massive systems, gluon loops or the smearing in transverse space. Although the implementation of the space–time picture in the simple string topologies is easy and straightforward, the picture gets much more intricate when more complicated topologies are addressed, for which the region offset has to be taken into account. Hence, most of the time and effort in this thesis has been spent in overcoming those issues, presented in section 4 and appendices A and B.

From the space–time implementation, the space–time hadron density has been addressed in terms of different measures. First, the correlation between the energy–momentum rapidity and the corresponding space–time quantity was presented, as well as the space–time longitudinal and transverse spectra, which are closely related to their energy–momentum counterparts. Then, the temporal and radial evolution of hadron production was tackled, both for pp and $q\overline{q}$ systems, illustrating, among other aspects, the different decays of primary and secondary hadrons and the fact that, at some point in time, the number of total hadrons in the system stabilizes, since all the hadrons left in the system are regarded as stable hadrons. Moreover, the fact that hadronization takes some time to start could also be deduced from the temporal evolution of pp and $q\overline{q}$ systems, leading to the study of the typical scales at which hadronization takes places.

The evolution of hadron production was further addressed by analysing the hadron density in the central region, showing that the outflux of hadrons from the central region of the z coordinate is greater than the influx. The first evidence of close-packing arises in those studies, illustrating that close-packing is important in pp collisions, especially close to the collision centre, where hadronization takes place.

The hadron density has been also tackled as a function of multiplicity for different models, with and without colour reconnection. The median radius at which hadrons are created significantly increases with multiplicity in the model with colour reconnection, while it remains almost fixed when colour reconnection is off. As a consequence, the hadron density increases dramatically when colour reconnection is off, while almost remaining constant with the default CR on scenario. Despite this, even when colour reconnection is taken into account, the particles are produced closely packed.

The close-packing of hadrons was further analysed by accounting for the number of hadrons overlapping with a newly generated one in the production rest frame, for different event multiplicities. In this case, the number of nearby hadrons increases with multiplicity,

implying that close-packing becomes increasingly important with multiplicity, presumably leading to unknown effects not addressed in Pythia. This hadronic overlap predominantly affects low-p_T hadrons, as observed in the study of the hadronic overlap as a function of the transverse momentum, implying that high-p_T jets do remain unaffected by the extreme close-packing experienced in low-p_T particle production. Therefore, analysing the possible effects of the hadronic overlap could be the starting point to address the unexpected behaviour of high-multiplicity pp collisions.

Even though different situations and details have been addressed in the space–time implementation, it can be improved in the future by including the exact location of each multiparton interaction and the offset given by parton showers. Moreover, secondary collisions among hadrons can also be included and studied in Pythia using the space–time implementation. Regarding the analysis of close-packing, its effects on the hadronization process have to be studied and understood. In fact, the density of hadron production could be the starting point of a new discussion on the formation of QGP, since it is related to the particle and energy density of the system. Analysing those aspects further could contribute to a better understanding of high-multiplicity pp collisions and, hence, of the formation of Quark-Gluon plasma.

A Space–time location of the final breakup

Hadronization is modelled in Pythia following a fragmentation process which iterates until the remaining invariant mass of the system is only sufficient to generate the last two hadrons (see section 4.1). At that point, a final breakup is generated between the two last consecutive breakups, one created stepping from the q side and the other from the $\overline{q}$ side. For the energy–momentum picture, the final breakup occurs in a fictitious final region, created from the combination of all the unused parts of all remaining regions.

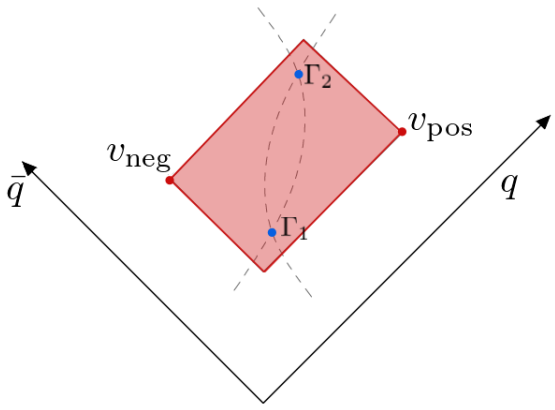


Figure 32: The two possibilities for the final breakup, in blue, and the final region, in red, in a $q\overline{q}$ system. The two old breakups, v_{neg} and v_{pos}, are also represented in red.

The procedure presented in section 3.4 and the process followed to generate the final breakup differ in the calculation of the longitudinal momentum of the hadron. This calculation is performed in the final region after the two hadron kinds have been established. Hence, the mass distribution of each of those hadrons is represented by hyperbolae, as explained in section 3.2.3. The two hyperbolae, corresponding to the mass constraint of each of the two final hadrons, might either cross in two different points or not cross at all. In the latter case, no solution can be found and the fragmentation process should be repeated again. The former case is illustrated for a $q\overline{q}$ system in Figure 32, where the blue dots represent the two points where the hyperbolae meet and the final region is depicted in red. Then, when the hyperbolae cross in two different locations, there are two possible options to describe the final breakup point. Since both possibilities have different Γ values, the probabilities for them to occur are given by eq. (3.17) (see section 3.2.3). For simplicity, eq. (3.17) is simplified in this specific case as $P(\Gamma) \approx \exp(-b\Gamma)$. Hence, the probability ratio of each of the two possibilities are

$$\begin{aligned} P(\Gamma_1) &= \frac{\exp(-b\Gamma_1)}{\exp(-b\Gamma_1) + \exp(-b\Gamma_2)} = \frac{1}{1 + \exp(-b(\Gamma_2 - \Gamma_1))} \\ P(\Gamma_2) &= \frac{\exp(-b\Gamma_2)}{\exp(-b\Gamma_1) + \exp(-b\Gamma_2)} = \frac{1}{1 + \exp(-b(\Gamma_1 - \Gamma_2))}. \end{aligned} \tag{A.54}$$

Then, the final breakup is obtained based on the probability ratios in the final region. The kinematics of the hadrons is also constructed in this final region.

Although the process presented so far is enough to develop the energy–momentum picture, it does not provide the information needed in the space–time picture. To implement the latter, the fractions $\hat{x}^{\pm}$ of each breakup with respect to the origin of the real region in which the breakup is located are needed. For the final breakup, those fractions are calculated in the final region. Since this region does not have a well-defined space–time offset (see section 4.1), a new method is needed to obtained the space–time location of this final breakup. The development of this method has been one of the main and time-consuming issues of this project. In particular, the major complication has been found in systems consisting of several gluons, specifically when the two old breakups are located in different regions. In those cases, knowing the region in which the final breakup is located is not possible. Since that aspect is essential to calculate the $\hat{x}^{\pm}$ fractions, several methods were tested to finally choose the ones presented in this appendix.

To explain the methods, we will refer to the old breakup closer to the q end as positive breakup, v_{pos}, the region at which it is located as positive region, and the final hadron generated from v_{pos} and the final breakup as positive hadron. Similarly, the negative breakup, v_{neg}, is the breakup closer to the $\overline{q}$ end, the negative region is the region in which it is located and the negative hadron, the one generated from the negative breakup and the final breakup.

The first implemented method consists in projecting the positive/negative hadron four-momentum on the positive/negative region. If the two old breakups are located in the

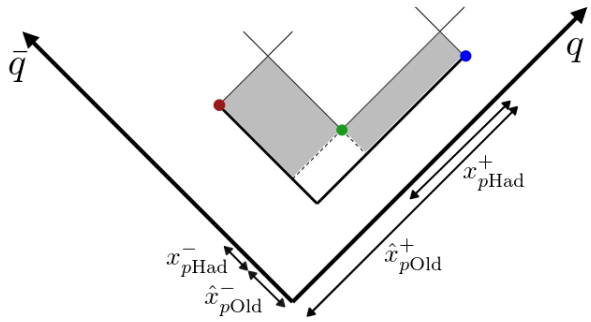


Figure 33: Final breakup point and final two hadrons, corresponding to the grey regions, in a $q\overline{q}$ system. The red and blue points are the previous breakups and the endpoints of the final region, while the green dots represents the final breakup.

same region, the positive and negative hadron four-momenta are projected on the same region. In that case, the final region is contained in the real region and they have the same transverse and longitudinal directions. This situation is exemplified in Figure 33, where the green point is the final breakup, the red and blue points are the old breakups, which correspond to the endpoints of the final region, and the grey squares represent the two final hadrons created. As can be seen in the figure, the $x^\pm$ fraction of the positive hadron are $x^\pm_{p\mathrm{Had}}$ and the $\hat{x}^\pm$ of the positive breakup are $\hat{x}^\pm_{p\mathrm{Old}}$. Then, the $\hat{x}^\pm_f$ of the final breakup can be obtained as

$$\begin{aligned} \hat{x}^+_f &= \hat{x}^+_{p\mathrm{Old}} - x^+_{p\mathrm{Had}} \\ \hat{x}^-_f &= \hat{x}^-_{p\mathrm{Old}} + x^-_{p\mathrm{Had}}. \end{aligned} \tag{A.55}$$

The same procedure can be followed with the negative breakup and the negative hadron. The expression can be derived as in the previous case

$$\begin{aligned} \hat{x}^+_f &= \hat{x}^+_{n\mathrm{Old}} + x^+_{n\mathrm{Had}} \\ \hat{x}^-_f &= \hat{x}^-_{n\mathrm{Old}} - x^-_{n\mathrm{Had}}. \end{aligned} \tag{A.56}$$

The solutions from eq. (A.55) and eq. (A.56) will agree only if the positive and the negative regions correspond to the same string region. The projection method can also be used when the positive and negative regions are different, in which case the longitudinal momentum of the positive or negative hadron is projected on the corresponding region, using either eq. (A.55) or eq. (A.56). If the final breakup is not located in that specific region, the values for $\hat{x}^\pm_f$ derived from this method do not satisfy the requirement $0 \leq \hat{x}^\pm_f \leq 1$. Since there is no possible way of finding the region in which the final breakup is located to use the projection method in that region, a new method is needed in such cases.

One of the main complication to obtain the $\hat{x}^\pm_f$ is the fact that the z value of the final breakup is not calculated, since it is not needed in the energy-momentum picture. If the projection method fails, the z value can be calculated from the Γ of the old breakups, the transverse masses of the final two hadrons, and the transverse mass of the final region.

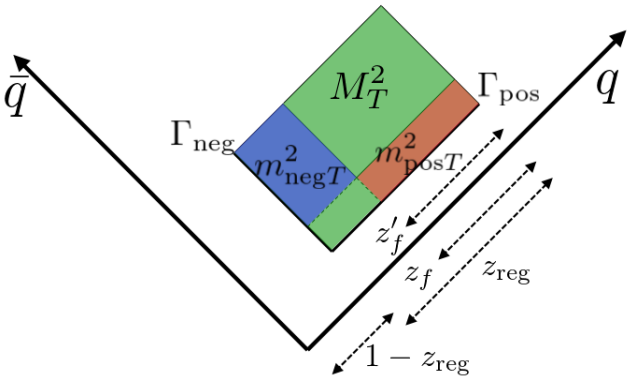


Figure 34: Final breakup and final two hadrons of a $q\bar{q}$ system. The blue and red areas represent the two final hadrons while the green area corresponds to the final region.

All these variables are depicted in Figure 34, along with z'_f and z_f which correspond to the z^+ fractions of the positive hadron in the final region and the real region, respectively. The $z_{\rm reg}$ variable in the figure represent the z^+ fraction taken from the negative breakup if the final breakup was not created. This value can be calculated from the relation give by eq. (3.16), which in this case gives

$$\Gamma_{\rm neg} = (1 - z_{\rm ref}) \left(\Gamma_{\rm pos} + \frac{m^2_{{\rm pos}T}}{z_{\rm reg}} \right) \tag{A.57}$$

where the final breakup was not taken into account. From this relation, the value of $z_{\rm reg}$ is found to be

$$z_{\rm reg} = \frac{\sqrt{\left(m^2_{{\rm pos}T} + \Gamma_{\rm neg} - \Gamma_{\rm pos}\right)^2 + 4m^2_{{\rm pos}T}\Gamma_{\rm pos}} - \left(m^2_{{\rm pos}T} + \Gamma_{\rm neg} - \Gamma_{\rm pos}\right)}{2\Gamma_{\rm pos}}. \tag{A.58}$$

Once the $z_{\rm reg}$ value is calculated, the $z_f = z_f^+$ can be determined, noting that $z_f = z'_f z_{\rm reg}$, which relates the positive hadron fractions in the final and the real regions. The same process can be followed to calculate z_f^-, using $m^2_{{\rm neg}T}$ instead of $m^2_{{\rm pos}T}$ and swapping the variables $\Gamma_{\rm pos}$ and $\Gamma_{\rm neg}$. Once the z_f^+ or z_f^- value is calculated, the region in which the final breakup is located can be deduced and the $\hat{x}^\pm$ fractions calculated.

Although the last procedure presented succeeds in the large majority of the cases, sometimes the region in which the final breakup is located cannot be found. In those few cases, the space–time location of the final breakup is calculated from the old breakups by determining the space–time location of the origin of the final region. As can be deduced from Figure 35, the space–time location of the CM of the final region is defined by $v_{\rm CM} = (v_{\rm pos} + v_{\rm neg})/2$. This final region can be treated as a $q\bar{q}$ system whose four-momentum is $K = k^+ + k^-$, where $k^\pm$ are the four-momenta vectors of the two endpoints. Following the same approach used to derive the early hadron production point (see section 4.1), the

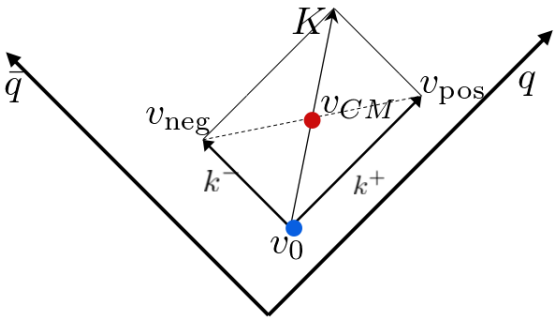


Figure 35: Final region of a $q\bar{q}$ system. The origin of the region is represented by v_0 while $v_{\rm CM}$ stands for the location of the CM of the region. The four-momentum of the final region is defined as K. As in other cases, $v_{\rm pos}$ and $v_{\rm neg}$ represent the old breakups.

space–time location of origin of the final region is given by

$$v_0 = v_{\rm CM} - \frac{1}{2}K. \tag{A.59}$$

Considering $\hat{x}_f'^{\pm}$ to be the $\hat{x}^{\pm}$ fractions of the final breakup in the final region, the space–time location of the final breakup can be calculated as

$$\begin{aligned} v_f &= v_0 + \hat{x}'^{+}k^+ + \hat{x}'^{-}k^- = \\ &= v_{\rm CM} + \left(\hat{x}'^{+} - \frac{1}{2}\right)k^+ + \left(\hat{x}'^{-} - \frac{1}{2}\right)k^-. \end{aligned} \tag{A.60}$$

Eq. (A.60) holds in the transverse rest frame, when the final region system is not evolving in time. That is not the case if the two old breakups are in different regions. In order to find an expression valid for all the cases, we define the variable $r = \sqrt{l^2/K^2}$, with $l^2 = -\left(v_{\rm neg} - v_{\rm pos}\right)^2$, to quantify how much the final string system differs from the system in the transverse rest frame. Then, the origin of the final region in the string system is determined by $v_0 = v_{\rm CM} - Kr/2$ and the general expression for the space–time location of the final breakup is

$$v_f = v_{\rm CM} + \left(\hat{x}'^{+} - \frac{r}{2}\right)k^+ + \left(\hat{x}'^{-} - \frac{r}{2}\right)k^- \tag{A.61}$$

The methods presented in this section are implemented in PYTHIA to obtain the space–time location of the final breakup. First the projection method is executed from the q or positive side. If it fails, the same method from the negative side is carried out. Whenever the projection method fails, the calculation of z^+ is performed as previously explained. In case of failure, the same method is carried out to calculate z^-. If none of the previous methods work, the space–time location of the final breakup is determined by eq. (A.61). The different procedures are carried out in order of accuracy, tested several times when deciding the methods that were going to be implement in PYTHIA.

B Correction to non-physical situations

As mentioned in section 3.2.3, by definition $0 < \hat{x}^{\pm} < 1$. Although this should always be the case, PYTHIA allows values outside that range whenever a step is taken from one region to a new region. Since the $\hat{x}^{\pm}$ fractions were only used to determine the energy and momentum of the hadrons, stepping outside the allowed range was not a problem, as long as the hadron energy was positive. Nevertheless, fractions outside the range $\hat{x}^{\pm}$ are a significant problem in the space–time picture, since they lead to negative times or negative squared invariant times of the breakup locations. In order to correct these unphysical situations, two corrections are applied in the space–time implementation. Both corrections are included before adding the region offset, the smearing in transverse space or the massive correction.

The first correction consists in adjusting the space–time location of the breakup with a fraction of the region four-momentum, requiring the new squared invariant time to be equal to zero, $\Gamma_{\text{new}} = 0$. The four-momentum of the region is defined by the four-momenta endpoint vectors, $p^{\pm}$, as $p_{\text{reg}} = p^{+} + p^{-}$. Then, the expression of the new breakup location is determined by

$$v_{\text{new}} = v_{\text{old}} + \xi p_{\text{reg}} \tag{B.62}$$

where ξ is a number, v_{old} and v_{new} are the same breakup space–time location before and after the correction, respectively. The value ξ is calculated by requiring $\Gamma_{\text{new}} = v_{\text{new}}^2 = 0$.

Another approach can be followed to correct the values of the breakup space–time location. In this case, the $\hat{x}^{\pm}$ fractions whose values are $\hat{x}^{\pm} > 1$ or $\hat{x}^{\pm} < 0$ are set equal to one and zero, respectively, and the space–time location is calculated again as presented in section 4.1 (eq. (4.31)).

The new breakup space–time location is determined following the two different approaches presented above. The option adopted is the one whose space–time components are closer to the space–time components of the breakup before the correction.

Acknowledgement

First of all, I would like to thank my supervisor, Torbjörn Sjöstrand, for his support, his patient, and for always being available to discuss all the possible questions that came to my mind. Thanks for guiding me throughout this project, for your time and your help.

Thanks to Anna, Virginia, Yosse and Konstantin for the fika breaks, the nights at Rydbergs and good moments that we have shared. You have always been there to cheer me up. Also to Marius, for his feedback in some parts of this thesis and for listening to me whenever I was stuck and needed some help. The past two years would have been really boring without all of you.

To Carolina and her family for making me feel part of your amazing family. You have always taken care of me and supported me. You know you will always be my Swedish-Venezuelan family.

Thanks to Daniel for all his support, for the countless Skype calls and flights, and for his help with some diagrams in this thesis. You are always there for me, no matter what, knowing exactly what to do to make me laugh.

Finally, a special thank you to my parents and my brother for being who they are. You have always supported me in every decision I have made, helping me to be who I wanted to be, specially when I got the crazy idea of becoming a physicist. This is the result, I hope you like it.

References

[1] M. Thomson, *Modern Particle Physis*, Cambridge University Press, 2013, ISBN 978-1-107-0342-6.

[2] F. Halzen and A. D. Martin *Quarks and Leptons: An Introductory Course in Modern Particle Physics*, John Wiley and Sons 1984, ISBN 0-471-88741-2.

[3] A. Buckley et al., *General purpose event-generators for LHC physics* Phys.Rept. 504 (2011) 145, doi: 10.1016/j.physrep.2011.03.005, arXiv:1101.2599 [hep-ph].

[4] M. H. Seymour and M. Marx, *Monte Carlo Event Generators* MCNET-13-05, arXiv:1304.6677 [hep-ph].

[5] T. Sjöstrand, S. Mrenna and P. Skands, PYTHIA *6.4 Physics and Manual* JHEP 0605 (2006) 026, doi: 10.1088/1126-6708/2006/05/026, arXiv:0603175 [hep-ph].

[6] T. Sjöstrand et al., *An Introduction to* PYTHIA *8.2*, Comput. Phys. Commun. 191 (2015) 159, doi: 10.1016/j.cpc.2015.01.024, arXiv:1410.3012 [hep-ph].

[7] M. Bahr et al., Eur. Phys. J. C58 (2008) 639-707, doi: 10.1140/epjc/s10052-008-0798-9, arXiv:0803.0883 [hep-ph].

[8] J. Bellm et al., *Herwig 7.0/Herwig++ 3.0 release note*, Eur. Phys. J. C76 (2016) no.4 196, doi: 10.1140/epjc/s10052-016-4018-8, arXiv:1512.01178 [hep-ph].

[9] T. Gleisberg et al., *Event generation with* SHERPA *1.1*, JHEP 090 (2009) 2007, doi: 10.1088/1126-6708/2009/02/007, arXiv:0811.4622 [hep-ph].

[10] B. Andersson, G. Gustafson, G. Ingelman and T. Sjöstrand, *Parton fragmentation and string dynamics*, Phys.Rept. 97, Nos 2 & 3 (1983) 31-145.

[11] T. Sjöstrand, *Jef fragmentation of multiparton configurations in a string framework*, Nuc. Phys. B248 (1984) 469-502.

[12] W. Busza, K. Rajagopal and W. van der Schee, *Heavy Ion Collisions: The Big Picture and the Big Questions*, Ann. Rev. Nucl. Part. Sci. 68 (2018) 1–49, arXiv:1802.04801 [hep-ph].

[13] ALICE Collaboration, J. Adam et al., *Enhanced production of multi-strange hadrons in high-multiplicity proton-proton collisions*, Nature Phys. 13 (2017) 535, doi: 10.1038/nphys4111, arXiv:1606.07424 [nucl-ex].

[14] CMS Collaboration, *Observation of long-range near-side angular correlation in proton-proton collisions at the LHC*, JHEP 09 (2010) 091, doi: 10.1007/JHEP09(2010)091, arXiv:1009.4122 [hep-ex].

[15] CMS Collaboration, *Measurement of long-range near-side two-particle angular correlations in pp collision at* $\sqrt{s} = 13$ *TeV*, Phys. Rev. Lett 116 (2016) 172302, doi: 10.1103/PhysRevLett.116.172302, arXiv:1510.03068 [nucl-ex]

[16] ATLAS Collaboration, *Observation of long-range elliptic azimuthal anisotropies in* $\sqrt{s} = 13$ *and 2.76 TeV pp Collisions with the ATLAS detector*, Phys. Rev. Lett. 116 (2016) 172301, doi: 10.1103/PhysRevLett.116.172301, arXiv:1509.04776 [hep-ex].

[17] S. Bethke et al., *Quantum Chromodynamics in Review of Particle Physics*, C. Patrignani et al. (Particle Data Group), Chin. Phys. C, 40, 100001 (2016) and 2017 update.

[18] T. Sjöstrand, *Monte Carlo Tools*, doi: 10.1201/b11865-14, arXiv:0911.5286 [hep-ph]

[19] B. Andersson, G. Gustafson and T. Sjöstrand, *Baryon production in jet fragmentation and* γ*-decay*, Physica Scripta 32 (1985) 574.

[20] P. Edén and G. Gustafson, *Baryon production in the string fragmentation picture*, Z.Phys. C75, (1997) 41.

[21] B. Andersson, G. Gustafson and B. Söderberg, *A general model for jet fragmentation*, Z. Phys. C20 (1983) 317, LU-TP-83-2.

[22] C. Patrignani *et al.* (Particle Data Group), Chin. Phys. C, 40, 100001 (2016) and 2017 update.

[23] P. Skands, S. Carrazza and J. Rojo, *Tuning* Pythia *8.1: the Monash 2013 Tune*, Eur. Phys. J. C74 (2014) no.8, 3024, doi: 10.1140/epjc/s10052-014-3024-y, arXiv:1404.5630 [hep-ph].