Evidence of Entanglement in Proton-Proton Collisions at the LHC

by
Alek Hutson

A dissertation submitted to the Department of Physics,

College of Natural Sciences and Mathematics

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in Physics

Chair of Committee: Rene Bellwied

Committee Member: Claudia Ratti

Committee Member: Andrew Renshaw

Committee Member: Andre Barato

Committee Member: Anna Vershynina

University of Houston
June 2024

ACKNOWLEDGMENTS

I'd like to express my sincere appreciation to the many people who have influenced my academic journey. In particular, I extend my deepest gratitude to my advisor, Dr. Rene Bellwied, whose guidance and influence have been paramount to the successful completion of this work and my growth as a scientist. I also want to thank the esteemed members of my dissertation committee: Dr. Claudia Ratti, Dr. Andrew Renshaw, Dr. Andre Barato, and Dr. Anna Vershynina. Their voluntary support and invaluable input have significantly enhanced my scientific approach in this research.

In addition, I wish to extend my heartfelt gratitude to the many friends and family members who have provided unwavering moral support throughout my academic journey. While the list of people who have supported me is long, I owe a great deal of my scientific identity to two family members in particular. My father, Jeremiah Hutson, instilled in me the importance of a strong work ethic and the relentless pursuit of one's aspirations. My grandfather, Lawrence Tune, ignited in me a profound and enduring curiosity for scientific knowledge from a young age.

My PhD journey could not have been successful without the influence of each person named here, and many more. I will continue to look toward the beacon of example provided by these people in my future pursuits as a scientist and as a person.

ABSTRACT

In an effort to better understand thermal behavior and particle yields in p-p collisions we recast the problem using the language of quantum information. In the last 50 years physicists have successfully used the parton model, to describe particle collisions. In the parton model the proton is put into a high momentum frame in which constituents are viewed as quasi-free. The proton wavefunction, described by quantum chromodynamics, exhibits a coherent superposition of quantum states and maintains unitary evolution, suggesting it is a pure quantum state. This pure state of quasi-free particles can be achieved through entanglement of the proton's constituents. We seek to show that this entanglement in the initial-state has a measurable effect on the evolution of the system and is the driving mechanism behind the thermal-like behavior and particle yields observed in the final-state. Recent theoretical predictions and experimental observations have demonstrated that entanglement in the initial state could survive in a strongly coupled system. Under this assumption we make a comparison between the distribution of information (parton number) in the initial state to the distribution of information (hadron number) in the final state. A comparison is also made between the entanglement entropy derived from the initial-state distribution and the thermodynamic-like Shannon entropy in the final-state distribution. Final-state distributions are extracted from experimental data collected using A Large Ion Collider Experiment (ALICE) at the Large Hadron Collider (LHC). In making this comparison between the initial and final state we observe a strong correspondence between the information in both states. On one hand, the comparison of moments calculated from an entangled Color Glass Condensate model with measured moments shows agreement in the spread of information between the two systems. On the other hand, calculations of entropy, which quantify the disorder of information, also show a consistent agreement. This correspondence in information spread and entropy gives a strong indication that entanglement survives the systems evolution and has a direct influence on the final state particle yields.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1

Background and Motivation

This first chapter is meant to establish the theoretical framework and motivation behind measuring information and entanglement in p-p collisions. In it we will describe the current understanding of proton structure and QCD evolution and how this understanding can be further developed using the language of quantum information. We will motivate this approach with a series of experimental results involving collision systems, strongly interacting matter, and entanglement.

1.1 Entanglement

The concept of quantum entanglement was first proposed by Albert Einstein, Boris Podolsky, and Nathan Rosen in their 1935 paper, often referred to as the EPR paradox [1]. This concept was further developed and named "entanglement" by Erwin Schrödinger in the same year [2]. Entanglement was a mathematical consequence of the non-separability of the wavefunction in composite quantum systems as described by the formalism of quantum mechanics. Einstein proposed the idea in the form of an apparent paradox described by a thought experiment:

In the thought experiment we consider two particles, each with a single degree of freedom exhibiting two possible states. The rules of quantum mechanics dictate that each of these particles exists in a superposition of the two states. The value of the state is deterministic only through an exchange of information in the form of a measurement. We then organize the wave functions of the particles such that they form a coherent state. In this state we can perform a measurement on one particle, collapsing the wavefunction, and immediately know the state of the other, thus collapsing its wavefunction instantaneously.

An instantaneous exchange of information between the two particles clashed with Einstein's theories of causality and locality. Classically, information is stored and transferred through mediums and fields associated with energy distributions, such as electromagnetic fields. All of these fields are subject to Einsteins rules stating that nothing can travel faster than the speed of light. But the EPR paradox showed an exchange of information infinitely faster than the speed of light.

Schrödinger further defined the concept of entanglement, emphasizing its fundamental role in quantum mechanics. He highlighted that entangled particles exhibit correlations that cannot be explained by classical physics. Schrödinger's work laid the groundwork for understanding the profound implications of entanglement, which later became a cornerstone of quantum information theory.

> "When two systems enter into temporary physical interaction due to known forces between them, and when after a time of mutual influence the systems separate again, then they can no longer be described in the same way as before. I would not call that one but rather the characteristic trait of quantum mechanics, the one that enforces its entire departure from classical lines of thought. By the interaction the two systems have become entangled."
>
> — Erwin Schrödinger [3]

1.1.1 A classic example

Perhaps the simplest way to understand quantum entanglement is through an example involving the the spin states of electrons. Spin is intrinsic angular momentum found in elementary particles which takes quantized values. Fermions are particles with half-integer spins, meaning that the magnitude of their spin in natural units is a half-integer, e.g., $\frac{1}{2}$, $\frac{3}{2}$, $\frac{5}{2}$, and so on. Electrons are an example of a fermion with spin equal to $\frac{1}{2}$. An electron's measured spin can take one of two distinct values: $\frac{1}{2}$, (which we call spin up) or $-\frac{1}{2}$, (spin down). A free electron exists in a superposition of these two degrees of freedom: spin up and spin down. All the information about the particle's spin state is contained in the wavefunction. In the formalism of quantum mechanics, i.e., bra-ket notation, we can represent this wavefunction of the spin Hilbert space as follows:

$$|\Psi\rangle = A|\uparrow\rangle + B|\downarrow\rangle \tag{1.1}$$

where A represents the probability amplitude of the particle being in a spin-up state once measured, and B represents the probability amplitude of the particle being in a spin-down state once measured. It is important to understand that the electron's state is simultaneously spin up and spin down before it is measured. This is proven in experiments like the Stern-Gerlach experiment [4].

Now let us imagine a state consisting of two electrons. This system of two particles would be represented by taking a tensor product of the two wavefunctions:

$$|\Psi_1\rangle \otimes |\Psi_2\rangle = |\Psi_1, \Psi_2\rangle = A|\uparrow,\uparrow\rangle + B|\uparrow,\downarrow\rangle + C|\downarrow,\uparrow\rangle + D|\downarrow,\downarrow\rangle \tag{1.2}$$

where A represents the probability amplitude of both particles being in a spin-up state, B represents the probability amplitude of the first particle being in a spin-up state and the

second being in a spin-down state, and so on. Up to now, we have used a notation commonly used in quantum mechanics called bra-ket notation, where each state is represented by a bra $\langle|$ or a ket $|\rangle$. We can also represent this sort of system in a matrix notation where states are represented by matrices whose elements are the probability amplitudes. This representation is far more useful when dealing with many particles and/or many-state systems. In the matrix representation, equation (2) would look like this:

$$|\Psi_1\rangle \otimes |\Psi_2\rangle = \rho = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \tag{1.3}$$

This represents a projection of the density matrix often denoted as ρ. The density matrix describes how the density of states is distributed in the given basis (in this case the spin state basis). In general, a density matrix can describe any type of state, including continuous states, and can include any number of particles. As the number of particles and the number of states each particle can take increases, the dimensionality of the matrix increases.[1] The off-diagonal elements of the matrix (in equation (3), these would be B and C) represent coherent or entangled states, meaning more than one quantum state is represented.

Going back to our previous example of two electrons, it is not hard to imagine a situation in which A and D are equal to 0. In many quantum systems the Pauli exclusion principle applies, which forbids two identical particles from occupying the same state. In this case, the density matrix becomes:

$$\Psi = \begin{pmatrix} 0 & B \\ C & 0 \end{pmatrix} \tag{1.5}$$

[1]The more relevant degrees of freedom (or basis) exist the more complex the density matrix becomes. The full density matrix is a convolution of the density matrices in each basis, for example if both spin and charge are relevant degrees of freedom:

$$\rho = \rho_{\text{spin}} \otimes \rho_{\text{charge}} \tag{1.4}$$

This is an example of a completely entangled state because we can no longer separate the wave function of these two particles into a tensor product of two distinct states. This means that the physical act of measuring one particle affects the probability distribution of the system. If I measure the state of one particle, I know the state of the other particle without measuring it because its spin orientation will be opposite.

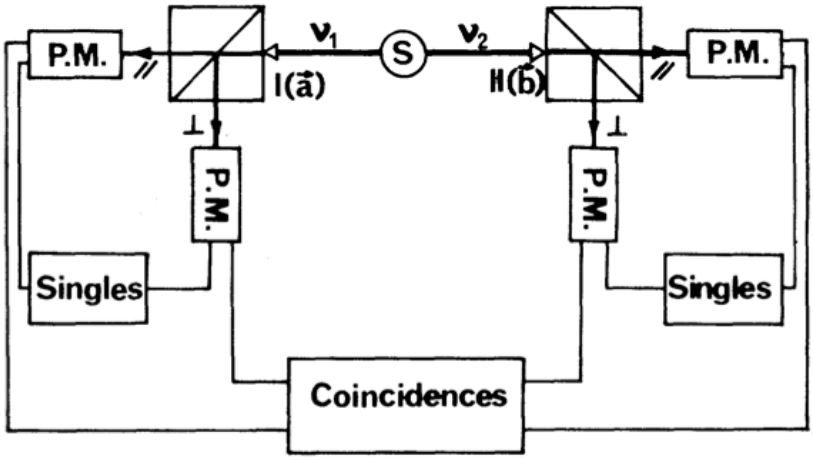

Figure 1.1: Diagram of the experimental setup Aspect and company used to make one of the first measurements of entanglement [5]

One of the most iconic experiments which verified this very concept was done by Alain Aspect, Philippe Grangier, and Gérard Roger in 1982 [5]. By measuring the polarization state of paired photons generated coherently (entangled) they demonstrated experimental evidence of entanglement.

1.1.2 Decoherence

Decoherence is a fundamental process in quantum mechanics that describes the loss of entanglement, or coherence, in a quantum system due to its interaction with the environment. This phenomenon explains the transition from quantum behavior to classical behavior, providing insight into the measurement problem and the apparent collapse of the wavefunction.

Mechanism of Decoherence

In quantum mechanics, a system is described by a density matrix ρ. When a system interacts with its environment, the total system-environment state can be described by the combined density matrix ρ_{SE}. The reduced density matrix of the system is obtained by tracing out the environmental degrees of freedom:

$$\rho_S = \mathrm{Tr}_E(\rho_{SE}). \tag{1.6}$$

If the initial state of the system is a pure state ψ_S, interaction with the environment typically leads to an entangled state between the system and the environment. The reduced density matrix ρ_S then represents a mixed state, reflecting the loss of coherence [6].

Decoherence Time

The timescale over which decoherence occurs is called the decoherence time, t_d. This timescale is often much shorter than the relaxation time, t_r, which is the timescale for energy exchange between the system and the environment. Decoherence time depends on various factors such as the strength of the system-environment interaction, the density of environmental states, and the temperature. A rough estimate of the decoherence time can be given by:

$$t_d \sim \frac{\hbar}{\lambda}, \tag{1.7}$$

where λ represents the interaction strength between the system and the environment [6]. [2]

[2] In our case of strongly interacting partons this interaction strength becomes extremely large making the decoherence time near 0. However, certain theoretical approaches show this proportionality may break down in some exotic states of matter, like color glass condensate discussed later in this chapter.

Decoherence in Quantum Systems

Decoherence plays a crucial role in various quantum systems and processes:

- **Quantum Computing:** Decoherence is a major challenge in maintaining quantum coherence in qubits, thereby affecting the reliability and performance of quantum computers.
- **Quantum Measurements:** Decoherence provides a mechanism for the apparent collapse of the wavefunction, as the system becomes entangled with the measuring apparatus and the environment, leading to classical outcomes.
- **Macroscopic Systems:** Decoherence explains why macroscopic objects do not exhibit quantum superpositions, as they are continuously interacting with their environments, leading to rapid decoherence.

Implications

Decoherence not only bridges the gap between quantum and classical worlds but also provides profound insights into the nature of quantum measurements. Decoherence involves a loss of information contained in the system to the environment. This information loss causes an entropy increase in the system and a local rise in temperature. Landauer quantified the lower limit on heat generated due to the loss of one bit of information as [7]:

$$Q = k_B T \ln 2, \tag{1.8}$$

Where Q represents the generated heat, k_B represents the Bolzmann constant, and T represents the temperature of the thermal reservoir. [3]

[3] In the context of hadronic interactions, which are purely quantum phenomenon, decoherence in the initial-state gives rise to the thermal-like behavior that follows. Quantifying the level of decoherence should allow us to extract a lower bound on the produced systems temperature.

1.1.3 Entanglement entropy

Entanglement entropy is a crucial concept in quantum information theory and quantum many-body physics. It quantifies the degree of entanglement between partitions of the whole. In a system of maximally entangled particles one can define an entropy between any two partitions of the Hilbert space and the entropies will be equal to one another, giving the entire system an entropy of 0. This is due to the fact that the system is pure and information is shared between all constituents. An interaction can occur in which a portion of the system interacts with the environment and loses coherence with the remaining elements of the system. Such an interaction will give rise to a non-zero entanglement entropy proportional to the loss in information between the system and the subsystem (or the information gained by the environment, depending on how one looks at it).

Von Neumann Entropy

The most commonly used measure of entanglement entropy is the von Neumann entropy. For a given pure state ρ:

$$\rho = \sum_i \lambda_i \rho_{A_i} \otimes \rho_{B_i}, \tag{1.9}$$

Where A and B represent two partitioned domains of the coherent state, and λ_i represent the Schmidt coefficients [8]. The reduced density matrix of subsystem A is obtained by tracing out the degrees of freedom of subsystem B: The von Neumann entropy S_A is defined as

$$S_A = -\mathrm{Tr}(\rho_A \log \rho_A). \tag{1.10}$$

This entropy quantifies the amount of quantum information shared between subsystems A and B. For the pure states, $S_A = S_B$ due to the symmetry of the bipartite system.

Properties of Entanglement Entropy

Entanglement entropy possesses several important properties:

- **Positivity:** $S_A \geq 0$.
- **Symmetry:** For a pure state ρ, $S_A = S_B$.
- **Subadditivity:** For a tripartite system A, B, and C, $S_{A\cup B} \leq S_A + S_B$.
- **Triangle Inequality:** $|S_A - S_B| \leq S_{A\cup B}$.

Entanglement in Quantum Field Theory

In quantum field theory (QFT), the concept of entanglement entropy can be extended to continuous systems. For a given region A in spacetime, the entanglement entropy is calculated by partitioning the Hilbert space into the degrees of freedom inside and outside A. The reduced density matrix ρ_A is obtained by tracing out the field degrees of freedom outside A. The resulting von Neumann entropy provides insights into the entanglement structure of the quantum fields.

Applications and Implications

Entanglement entropy has profound implications in various areas of physics:

- **Quantum Information:** It quantifies the resources required for quantum communication and computation [7].
- **Condensed Matter Physics:** Entanglement entropy helps characterize different phases of matter, especially in topological phases and critical systems [9].

- **Quantum Gravity:** The holographic principle and the AdS/CFT correspondence relate the entanglement entropy of a conformal field theory to the geometry of a dual gravitational theory [10].

Understanding entanglement entropy is fundamental for exploring the intricate connections between quantum mechanics, information theory, and the geometry of spacetime.

This work will provide a unique example showing how entanglement entropy, generated in the early stages of particle collisions, maps to a thermodynamic-like entropy in the final measured state.

1.2 Standard Model

The Standard Model of particle physics is a comprehensive theory that describes the fundamental particles and their interactions, except for gravity. It successfully unifies the electromagnetic, weak, and strong forces within a single framework, providing an accurate description of a wide range of experimental phenomena. This section provides a detailed overview of the Standard Model, covering its theoretical foundations, particle content, interactions, and experimental validation.

1.2.1 Theoretical Foundations

The Standard Model is based on the principles of quantum field theory and gauge invariance. It is a gauge theory with the gauge group:

$$SU(3)_C \times SU(2)_L \times U(1)_Y, \tag{1.11}$$

where $SU(3)_C$ represents the strong interaction (quantum chromodynamics, QCD), $SU(2)_L$ represents the weak interaction, and $U(1)_Y$ represents the hypercharge associated with the

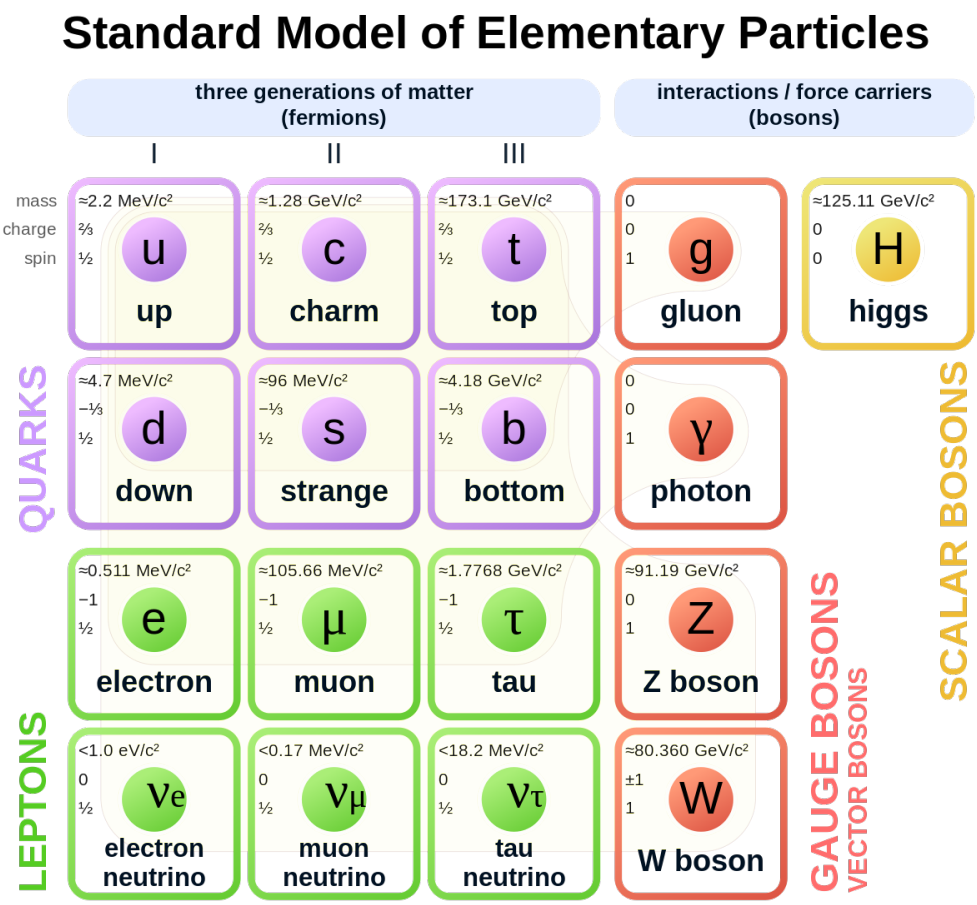

Figure 1.2: The Standard Model of Particle physics [11]

electromagnetic interaction [12]. In a gauge theory, symmetry groups define the invariances under certain transformations that the fields of the theory must respect. These groups govern the interactions between particles, ensure conservation laws, and define the fundamental forces observed in nature.

Gauge Symmetry and the Lagrangian

The Standard Model Lagrangian can be divided into several parts described below [13]:

Gauge Field Terms The gauge field terms describe the dynamics of the gauge bosons associated with the the strong interaction (gluons), weak interaction (W and Z bosons), and the electromagnetic interaction (photons):

$$\mathcal{L}_{\text{gauge}} = -\frac{1}{4}G^a_{\mu\nu}G^{a\mu\nu} - \frac{1}{4}W^i_{\mu\nu}W^{i\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu}, \tag{1.12}$$

where $G^a_{\mu\nu}$, $W^i_{\mu\nu}$, and $B_{\mu\nu}$ are the field strength tensors for the $SU(3)_C$, $SU(2)_L$, and $U(1)_Y$ gauge fields, respectively.

Fermion Terms The fermion terms describe the kinetic and mass terms for the quarks and leptons:

$$\mathcal{L}_{\text{fermion}} = \sum_f \bar{\psi}_f(i\gamma^\mu D_\mu - m_f)\psi_f, \tag{1.13}$$

where ψ_f represents the fermion fields, γ^μ are the Dirac matrices, D_μ is the covariant derivative, and m_f are the fermion masses.

Higgs Field Terms The Higgs field gives mass to the fermions and W and Z bosons. Its terms describe the dynamics of the Higgs boson and its interactions with other fields:

$$\mathcal{L}_{\text{Higgs}} = (D^\mu\phi)^\dagger(D_\mu\phi) - V(\phi), \tag{1.14}$$

where ϕ is the Higgs field, and $V(\phi)$ is the Higgs potential, given by:

$$V(\phi) = \mu^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2. \tag{1.15}$$

Interaction Terms The interaction terms describe the couplings between the fermions, gauge bosons, and the Higgs field. For example, the Yukawa interactions between the Higgs

field and the fermions are given by:

$$\mathcal{L}_{\text{Yukawa}} = -\sum_f y_f \bar{\psi}_f \phi \psi_f, \tag{1.16}$$

where y_f are the Yukawa coupling constants [14].

1.2.2 Particle Content

The Standard Model includes a variety of fundamental particles, categorized into three main groups: gauge (vector) bosons, fermions, and the Higgs (scalar) boson.

Gauge Bosons

Gauge bosons are vector bosons that act as the force carriers of the Standard Model. They mediate the fundamental interactions between the particles:

- **Photon** (γ): Mediates the electromagnetic interaction.
- **W and Z Bosons** ($W^{\pm}$, Z): Mediate the weak interaction.
- **Gluons** (g): Mediate the strong interaction. There are eight gluons corresponding to the $SU(3)_C$ gauge group.

Fermions

Fermions are the matter particles of the Standard Model. They are divided into quarks and leptons, each coming in three generations:

- **Quarks:**
  - First Generation: Up (u), Down (d)

  - Second Generation: Charm (c), Strange (s)
  - Third Generation: Top (t), Bottom (b)

- **Leptons:**
  - First Generation: Electron (e), Electron Neutrino (ν_e)
  - Second Generation: Muon (μ), Muon Neutrino (ν_μ)
  - Third Generation: Tau (τ), Tau Neutrino (ν_τ)

Each fermion has a corresponding antiparticle with opposite charge.

Higgs Boson

Higgs boson is a scalar particle responsible for giving mass to the W and Z bosons and fermions through the Higgs mechanism. Its discovery at the LHC in 2012 was a significant milestone in confirming the Standard Model [15].

1.2.3 Experimental Validation

The Standard Model has been extensively validated through numerous experiments. Key achievements include the discovery of the Higgs boson, precision measurements of the properties of the gauge bosons, and the verification of the predicted particle interactions.

Collider Experiments

High-energy collider experiments, such as those conducted at the LHC, have been crucial in testing the Standard Model. The discovery of the Higgs boson and the measurement of its properties are among the most significant results. Additionally, precise measurements of the electroweak parameters, such as the masses and couplings of the W and Z bosons, provide stringent tests of the Standard Model [16].

Neutrino Experiments

The standard model while powerful is far from complete. Aside from failing to describe phenomena like gravity, and dark matter, the study of neutrino oscillations contradicts the original formulation of the standard model. Neutrino oscillation experiments have provided evidence for neutrino masses and mixing, which are not accounted for in the Standard Model [17]. These results suggest the need for an extension of the Standard Model.

1.3 Proton structure

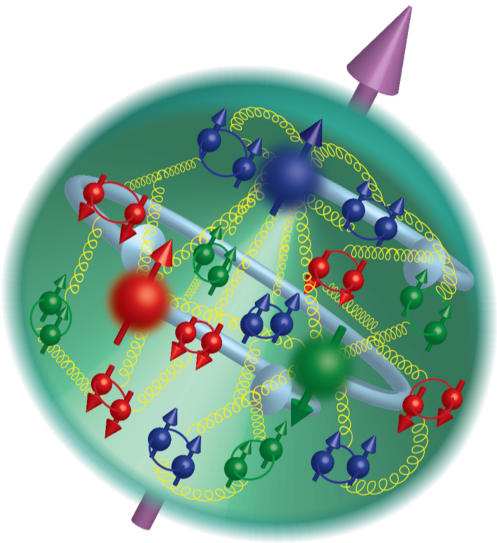

Figure 1.3: Argonne National Laboratory. "3D Structure of Protons and Neutrons." Argonne National Laboratory, U.S. Department of Energy, n.d., https://www.anl.gov/phy/3d-structure-of-protons-and-neutrons.

The proton was discovered in 1917 by Ernest Rutherford [18]. During that time the proton was thought to be an elementary particle. This misconception was understandable given that the proton is stable, and exhibits unitary charge and 1/2 spin in natural units, making it a pure quantum state. It was later discovered that the proton was not elementary

but did in fact have a complex substructure described by the parton model. This complex internal structure can be probed through various high-energy experiments, providing insights into Quantum Chromodynamics (QCD) and the behavior of strongly interacting particles.

1.3.1 Parton Model

The parton model was first proposed by Richard Feynman in a series of lectures and subsequently in a published paper in 1969 [19]. This model places the proton in a high momentum frame in which it is Lorentz contracted into a pancake-like object. In this frame, the proton can be described by a collection of point-like constituents called partons. These partons include quarks and gluons, which carry the momentum and quantum numbers of the proton. The parton model successfully explained the scaling behavior observed in deep inelastic scattering experiments, where electrons were scattered off protons, revealing the substructure within the proton. This was a significant step towards the development of Quantum Chromodynamics (QCD), the theory describing the strong interactions between quarks and gluons. The pioneering experiments by Friedman, Kendall, and Taylor provided the first experimental evidence for the parton substructure of protons [20, 21], fundamentally changing our understanding of particle physics and laying the groundwork for future discoveries in the field.

1.3.2 Parton Content

The distribution of partons which influence the protons wavefunction is dependent on the energy scale (Q^2) at which the proton is probed and is typically given as a function on the relative momentum carried by each parton (x).

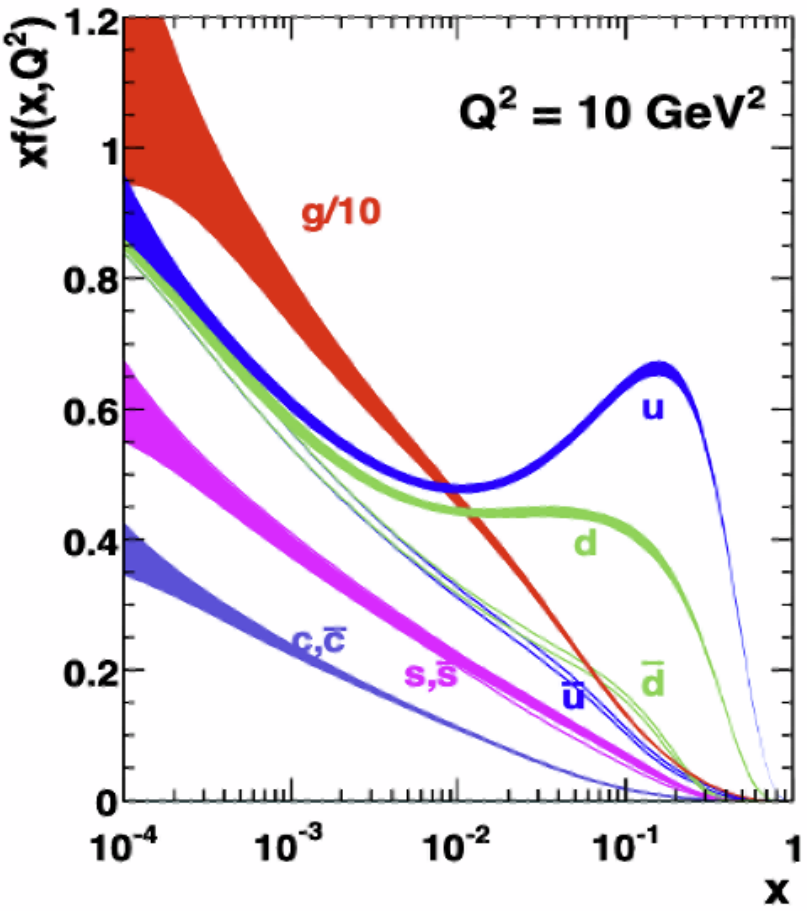

Figure 1.4: Distribution of each parton contribution to the protons wave function at $Q^2 = 10$ GeV^2 according to [22]

Valence Quark Contribution

Valence quarks are the primary constituents of the proton and contribute significantly to its quantum numbers, such as charge and baryon number. The proton consists of two up quarks and one down quark, giving it a net charge of +1. While valence quarks contribute to the proton's spin, their spins alone do not account for the total spin of 1/2. The spin structure of the proton is more complex, involving not just the valence quarks but also contributions from sea quarks and gluons. The dynamics and interactions among these components contribute to the overall spin of the proton through both intrinsic spin and

orbital angular momentum.

Sea Quark Contribution

Sea quarks, which are transient quark-antiquark pairs generated by quantum fluctuations, play a crucial role in the proton's internal dynamics. While they do not define the proton's primary quantum numbers, they contribute to its mass and influence its spin structure. The contribution of sea quarks to the proton's spin is complex, as they participate in the intricate interactions mediated by gluons. These interactions add to the overall spin and momentum distributions within the proton, revealing the deep and dynamic structure of quantum chromodynamics.

Gluon Contribution

Gluons, the carriers of the strong force, are vital in binding quarks together within the proton. They are massless and electrically neutral but play a significant role in the proton's mass and spin. Gluons contribute to the proton's spin through their intrinsic spin and the orbital angular momentum of the quarks and gluons. This contribution is crucial, as gluons mediate the interactions that hold the proton together and thus significantly impact its internal dynamics and observable properties.

1.4 Quantum Chromodynamics

Quantum Chromodynamics (QCD) is the theory of the strong interaction, one of the four fundamental forces in nature. It describes the interactions between quarks and gluons, the fundamental constituents of hadrons, such as protons and neutrons. QCD is a non-Abelian gauge theory based on the symmetry group $SU(3)_C$, where the subscript C stands for color charge, the equivalent of electric charge in Quantum Electrodynamics (QED).

1.4.1 The QCD Lagrangian

The QCD Lagrangian describes the dynamics of quarks and gluons. It is given by:

$$\mathcal{L}_{\text{QCD}} = \sum_f \bar{\psi}_f (i\gamma^\mu D_\mu - m_f)\psi_f - \frac{1}{4} G^a_{\mu\nu} G^{a\mu\nu}, \tag{1.17}$$

where ψ_f represents the quark field of flavor f, m_f is the quark mass, D_μ is the covariant derivative, and $G^a_{\mu\nu}$ is the gluon field strength tensor. The covariant derivative D_μ and the gluon field strength tensor $G^a_{\mu\nu}$ are defined as:

$$D_\mu = \partial_\mu - i\alpha_s T^a A^a_\mu, \tag{1.18}$$

$$G^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + \alpha_s f^{abc} A^b_\mu A^c_\nu, \tag{1.19}$$

where α_s is the QCD coupling constant, T^a are the generators of $SU(3)_C$, A^a_μ are the gluon fields, and f^{abc} are the structure constants of $SU(3)_C$ [12].

1.4.2 Asymptotic Freedom and Confinement

QCD exhibits two remarkable properties: asymptotic freedom and confinement. Asymptotic freedom implies that at high energies (short distances), the interaction between quarks becomes weaker, allowing quarks to behave as nearly free particles. This was discovered by Gross, Wilczek, and Politzer in 1973 and is a key feature of QCD [23, 24]. The running of the strong coupling constant $\alpha_s(Q^2)$ with the energy scale Q^2 is given by [25]:

$$\alpha_s(Q^2) = \frac{4\pi}{\beta_0 \ln(Q^2/\Lambda^2)} \tag{1.20}$$

where β_0 is the beta function coefficient and Λ is the QCD scale parameter.

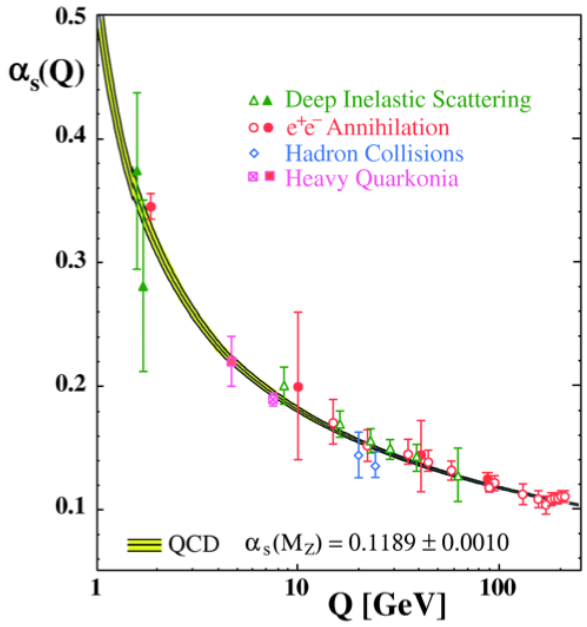

Figure 1.5: Coupling constant as a function of the energy scale Q [25]. Averaging over each of the results gives a baseline coupling constant at the energy scale of the Z boson mass equal to about 0.1189

Confinement, on the other hand, means that quarks and gluons cannot exist as free particles at low energies (large distances). They are permanently confined within hadrons. A consequence of confinement in QCD is that particles form color-neutral states. Overcoming this confinement requires high energy densities, but this system is short-lived, making the direct detection of colored objects impossible within our current experimental capabilities. This property is not yet fully understood analytically, but it is supported by lattice QCD calculations and experimental observations [26].

1.4.3 Phases of QCD Matter

QCD predicts different phases of matter depending on temperature and baryon density. The QCD phase diagram, which maps the different phases of QCD matter, is an active area of research.

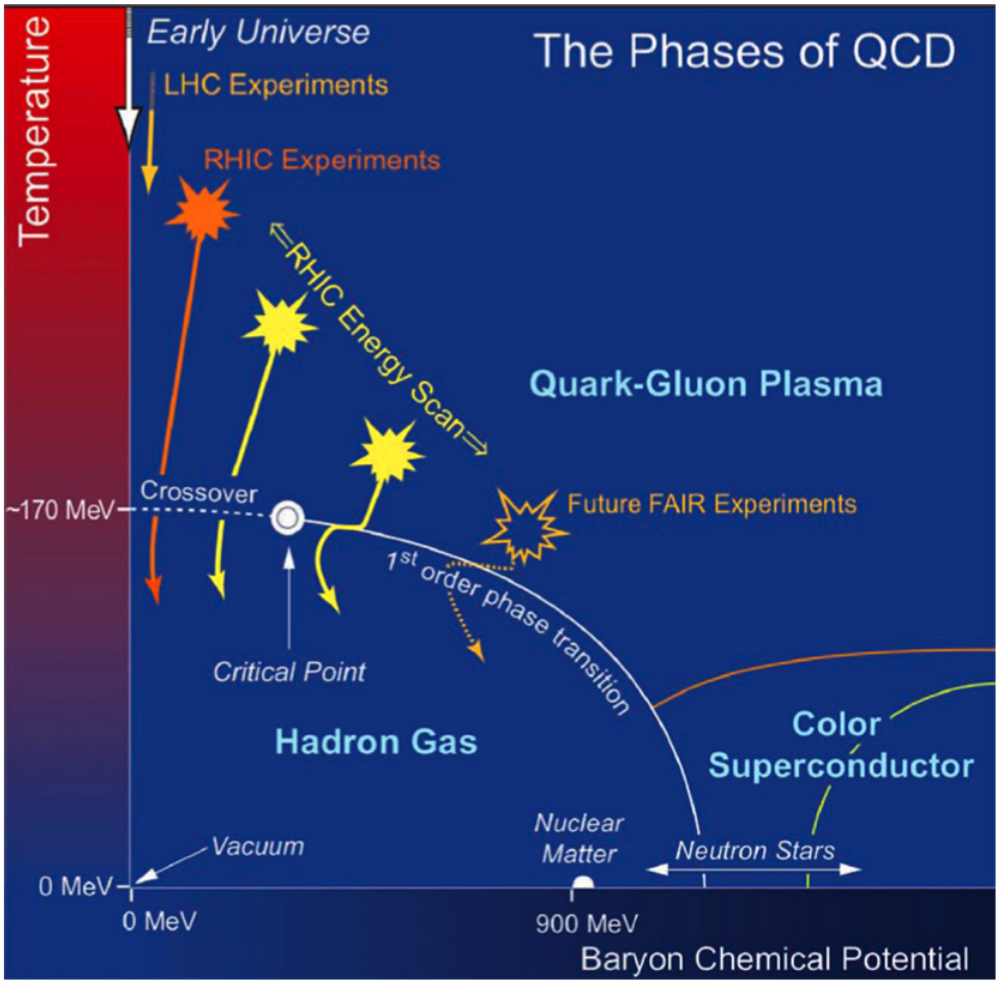

Figure 1.6: Phases diagram of QCD matter in the baryonic chemical potential and temperature plane [27].

At high temperatures or densities, quarks and gluons are expected to exist in a deconfined state known as the Quark-Gluon Plasma (QGP). The transition can be thought of as a melting of hadronic matter (like protons) into a fluid like state. This state of matter is believed to have existed in the early universe shortly after the Big Bang and can be recreated in heavy-ion collisions at facilities like the Large Hadron Collider (LHC) and the Relativistic Heavy Ion Collider (RHIC).

The phase transition from hadronic matter to QGP is studied using lattice QCD simulations. The transition temperature is found to be $T_c \approx 158$ MeV at $\mu_B = 0$ [28]. This has

been experimentally studied in high energy hadronic collisions.

We hope to better understand the behavior of entanglement in these different phases of matter, and across phase transitions where there is a defined change in the degrees of freedom.

1.4.4 Matter Generation in Scattering Processes

We understand from Einsteins famous equation $E = mc^2$ that energy can be converted into matter [29]. Collision systems behave as matter factories converting energy from the collision into matter in the form of particles. Quantum Chromodynamics (QCD) plays a crucial role in understanding particle production in both hadronic interactions and interactions such as those seen in Deep Inelastic Scattering (DIS). These processes provide critical insights into the dynamics of quarks and gluons and the structure of hadrons. Cross sections are fundamental quantities in these studies, representing the probability that a particular interaction or scattering process will occur.

Hadronic Collisions

In hadronic collisions, such as proton-proton, proton-nucleus, and nucleus-nucleus interactions, QCD describes the production of particles through the interactions of partons (quarks and gluons) within the colliding hadrons. The cross sections for these processes are calculated using perturbative QCD (pQCD) at high energies, where the strong coupling constant α_s is small, allowing for a controlled expansion. The inclusive cross section for producing a particle h in a hadronic collision can be written as [30]:

$$d\sigma(pp \to h + X) = \sum_{a,b} \int dx_a \int dx_b \, f_{a/p}(x_a, Q^2) \, f_{b/p}(x_b, Q^2) \, d\hat{\sigma}(ab \to h + X), \qquad (1.21)$$

where $f_{a/p}(x_a, Q^2)$ and $f_{b/p}(x_b, Q^2)$ are parton distribution functions (PDFs) of the partons a and b, and $d\hat{\sigma}(ab \to h + X)$ is the partonic cross section.

Deep Inelastic Scattering

In Deep Inelastic Scattering (DIS), high-energy electrons (or other leptons) scatter off nucleons, resulting in the exchange of a virtual photon, γ^*, with large momentum transfer. This process probes the internal structure of protons and neutrons, revealing information about the distribution and dynamics of quarks and gluons. The inclusive cross section for DIS can be expressed similarly to hadronic collisions. The differential cross section is given by [31]:

$$\frac{d^2\sigma}{dxdQ^2} = \frac{4\pi\alpha^2}{xQ^4}\left[\left(1 - y + \frac{y^2}{2}\right)F_2(x, Q^2) - \frac{y^2}{2}F_L(x, Q^2)\right], \tag{1.22}$$

where $F_2(x, Q^2)$ and $F_L(x, Q^2)$ are the structure functions, and $y = \frac{\nu}{E}$ is the inelasticity variable.[4]

Setting the Stage for Further Studies

Understanding particle production in hadronic interactions and DIS is fundamental for exploring the structure of matter at the smallest scales. QCD provides the theoretical framework to describe these processes, allowing us to extract parton distribution functions and study the behavior of quarks and gluons under various conditions. The insights gained from these studies pave the way for further exploration of hadronic collisions and the properties of QCD matter in extreme environments.

In the context of entanglement in p-p collisions the QCD framework helps to define the initial distribution of information inside the proton, and how energy scaling is related to

[4]PDFs and structure functions are related to the distribution of partons inside the proton, this will be further defined and discussed in the next chapter.

that distribution. This initial state information is extracted from DIS cross sections and compared to data from hadronic collisions.

1.5 Deep Inelastic Scattering

Deep inelastic scattering (DIS) is a powerful tool used to probe the internal structure of protons and neutrons, providing crucial insights into the partonic constituents, such as quarks and gluons. This process involves high-energy electrons (or other leptons) scattering off nucleons, resulting in the exchange of a virtual photon, γ^*, with large momentum transfer. Employing Quantum Electrodynamics (QED) probes in DIS experiments enables the disentangling of QCD effects from electroweak interactions, thus offering a precise method to investigate the non-perturbative aspects of QCD, such as the distribution of partons and their dynamics within nucleons.

1.5.1 Kinematics of Deep Inelastic Scattering

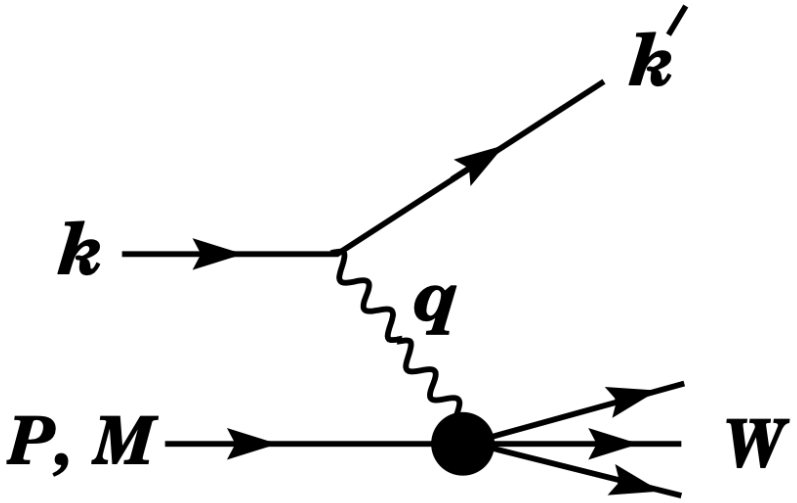

Figure 1.7: Feynmann Diagram representing the scattering of an electron off a proton [31].

The fundamental variables describing DIS are the Bjorken scaling variable x and the four-momentum transfer squared Q^2. Bjorken-x represents the momentum fraction of a struck parton, and Q^2 represents the energy scale of the interaction. Consider an electron with initial four-momentum $k = (E, \vec{k})$ scattering off a proton with four-momentum $P = (E_P, \vec{P})$. The electron emerges with final four-momentum $k' = (E', \vec{k'})$. The four-momentum of the exchanged virtual photon is $q = k - k'$.

The invariant mass squared of the hadronic system W^2 is given by:

$$W^2 = (P + q)^2. \tag{1.23}$$

Using the relations $q^2 = -Q^2$ and $Q^2 = -(k-k')^2$, we can define the Bjorken scaling variable x as:

$$x = \frac{Q^2}{2P \cdot q}. \tag{1.24}$$

In the laboratory frame, where the proton is initially at rest, this reduces to:

$$x = \frac{Q^2}{2M\nu}, \tag{1.25}$$

where M is the proton mass and $\nu = E - E'$ is the energy transferred to the proton.

The variable Q^2 represents the negative of the four-momentum transfer squared and is given by:

$$Q^2 = -q^2 = -(k - k')^2 = 4EE' \sin^2\left(\frac{\theta}{2}\right), \tag{1.26}$$

where θ is the scattering angle of the electron [31].

1.5.2 Structure Functions

In DIS, the cross section for the scattering process is related to the structure functions $F_1(x, Q^2)$ and $F_2(x, Q^2)$, which encode information about the distribution of quarks within the nucleon. The differential cross section for unpolarized electrons scattering off unpolarized protons is given by [31]:

$$\frac{d^2\sigma}{d\Omega dE'} = \frac{\alpha^2}{Q^4}\frac{E'}{E}L_{\mu\nu}W^{\mu\nu}, \tag{1.27}$$

where α is the fine-structure constant, $L_{\mu\nu}$ is the leptonic tensor, and $W^{\mu\nu}$ is the hadronic tensor.

The hadronic tensor $W^{\mu\nu}$ can be expressed in terms of the structure functions:

$$W^{\mu\nu} = \left(-g^{\mu\nu} + \frac{q^\mu q^\nu}{q^2}\right) F_1(x, Q^2) + \left(P^\mu - \frac{P\cdot q}{q^2}q^\mu\right)\left(P^\nu - \frac{P\cdot q}{q^2}q^\nu\right)\frac{2x}{Q^2}F_2(x, Q^2). \tag{1.28}$$

1.5.3 Parton Distribution Functions and Cross Sections

The parton distribution functions (PDFs) describe the probability densities of finding a quark or gluon with a given momentum fraction x inside the proton. These PDFs are extracted from experimental measurements of cross sections in DIS and other high-energy processes.

The inclusive DIS cross section can be written as:

$$\frac{d^2\sigma}{dxdQ^2} = \frac{4\pi\alpha^2}{xQ^4}\left[\left(1 - y + \frac{y^2}{2}\right)F_2(x, Q^2) - \frac{y^2}{2}F_L(x, Q^2)\right], \tag{1.29}$$

where $y = \frac{\nu}{E}$ is the inelasticity variable, and $F_L(x, Q^2)$ is the longitudinal structure function. In the scaling limit (high Q^2 and fixed x), the structure function $F_2(x, Q^2)$ is directly related

to the PDFs through a convolution integral:

$$F_i = \sum_a C_i^a \otimes f_a, \tag{1.30}$$

where $\otimes$ denotes the convolution integral:

$$C \otimes f = \int_x^1 \frac{dy}{y} C(y) f\left(\frac{x}{y}\right), \tag{1.31}$$

and where the coefficient functions C_i^a are given as a power series in α_s. The parton distribution f_a corresponds, at a given x, to the density of parton a in the proton.

The process of extracting PDFs involves fitting theoretical predictions to the measured cross sections. The PDFs are usually parameterized at an initial scale Q_0^2 and evolved to higher Q^2 using the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) equations.[5]

1.5.4 Experimental Measurements of Cross Sections

Experimentally, the measurement of DIS cross sections requires precise determination of the scattered electron's energy and angle, as well as the initial energy of the electron beam. The differential cross sections are measured as functions of x and Q^2, and integrated over the kinematic ranges accessible to the experiment.

Modern DIS experiments, such as those conducted at the HERA collider, have provided a wealth of data on the structure functions and PDFs [32]. The analysis of these data involves correcting for detector effects, radiative corrections, and other systematic uncertainties.

The extraction of PDFs from the measured cross sections involves sophisticated global fits, taking into account data from various processes and experiments. These global fits

[5]How DGLAP equations are used to extrapolate PDFs is described in a later chapter defining the initial state

provide a comprehensive picture of the parton distributions and their uncertainties.

1.6 Hadronic Collisions

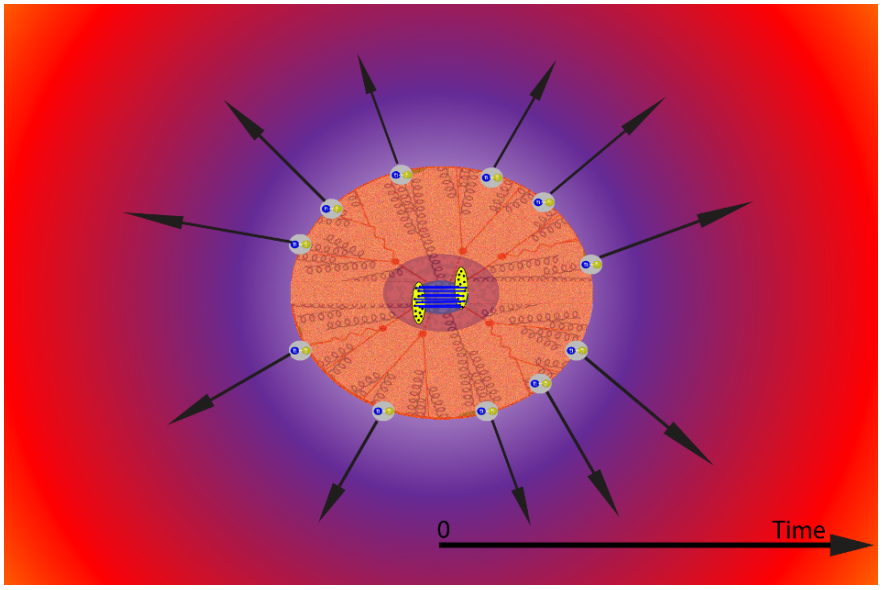

Figure 1.8: Visual representation of the evolution of a hadronic collision. Starting from a state of thermalized partons the system melts into a QGP which expands until the energy density is such that QCD confinement is reestablished through the formation of hadrons.

Hadronic collisions, such as those involving protons, neutrons, or nuclei, are fundamental processes studied in high-energy physics to understand the behavior of matter under extreme conditions. These collisions can lead to the formation of exotic states of matter, such as the Quark-Gluon Plasma (QGP), and provide insights into the mechanisms of hadronization, where quarks and gluons combine to form hadrons. This section explores the dynamics of hadronic collisions, the formation of QGP, the process of hadronization, and the signs of collectivity observed in high-energy proton-proton (p-p) collisions.

1.6.1 Quark-Gluon Plasma (QGP)

The Quark-Gluon Plasma (QGP) is a state of matter in which quarks and gluons, normally confined within hadrons, are free to move within a hot, dense medium. This state is believed to have existed in the early universe shortly after the Big Bang and can be recreated in high-energy heavy ion collisions at facilities such as the Large Hadron Collider (LHC) and the Relativistic Heavy Ion Collider (RHIC).

Formation of QGP

The formation of QGP occurs in high-energy collisions when the energy density exceeds a critical value, leading to deconfinement of quarks and gluons. The conditions necessary for QGP formation are typically achieved in nucleus-nucleus collisions, such as those involving gold (Au) or lead (Pb) nuclei. However, recent experimental evidence suggest the possible formation of QGP in small systems like p-p collisions at high energies.[6]

After the initial collision a large amount of energy is deposited in a small volume, creating a hot and dense medium. This medium undergoes rapid thermalization (1 fm/c) [7], resulting in the formation of QGP. The QGP then expands and cools, eventually undergoing a phase transition back to hadronic matter.

Properties of QGP

QGP exhibits several unique properties, including:

- **High Temperature and Density:** The temperature of QGP can reach several trillion Kelvin, and the energy density can be several times that of normal nuclear matter.

[6] see next section 1.6.2

[7] This rapid thermalization can be described using entanglement. This will be expanded upon in later sections

- **Collective Behavior:** QGP behaves like a near-perfect fluid with very low viscosity, leading to strong collective flow patterns.

- **Color Screening:** The color charges of quarks and gluons are screened at short distances, reducing the strength of the strong force between them.

The study of QGP properties is crucial for understanding the strong interaction and the behavior of matter under extreme conditions.

Typically in hot-dense systems, such as the QGP, decoherence occurs and evidence of entanglement is lost. Recent theoretical predictions suggest that entanglement may be able to survive in a strongly coupled system [33], this work hopes to provide some preliminary experimental evidence of this behavior.

1.6.2 Signs of Collectivity in High-Energy Proton-Proton (p-p) Collisions

Traditionally, QGP formation was expected only in heavy ion collisions due to the larger volume and higher energy densities. Recent experimental results have shown signs of collectivity in high-energy proton-proton (p-p) collisions, indicating the possible formation of QGP even in these smaller systems.

As a first study in searching for signs that entanglement might survive the QGP and the proceeding phase transitions we will look at these small systems (p-p) in which signs of collectivity have been observed.

Elliptic and Higher-Order Flow

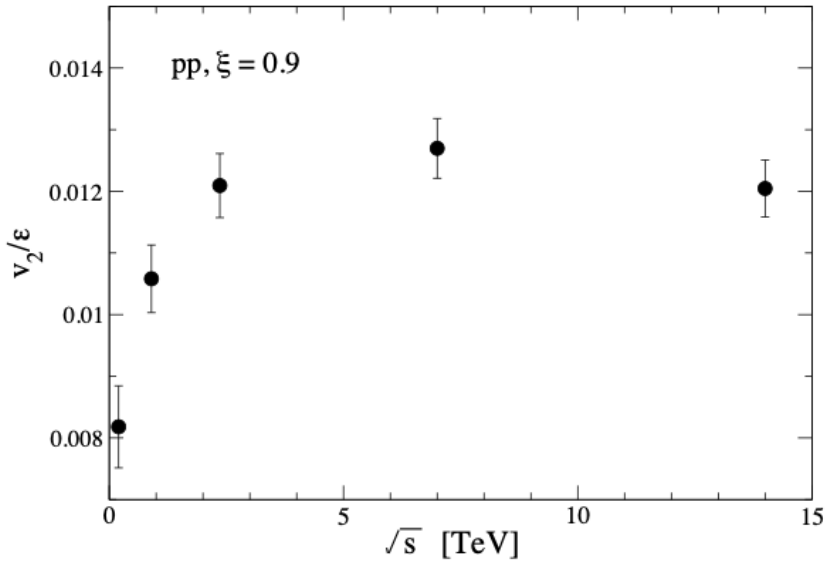

Figure 1.9: Scaled elliptic flow as a function of collision energy for min bias pp collisions at LHC energies [34].

One of the key signatures of collectivity is the observation of elliptic flow (v_2) and higher-order flow coefficients (v_3, v_4, etc.) [35].

$$v_n = \langle \cos[n(\phi - \Psi_n)] \rangle, \tag{1.32}$$

where n is the order of the flow coefficient, ϕ is the azimuthal angle of the emitted particle, and Ψ_n is the azimuthal angle of the n-th order event plane. These flow coefficients describe the azimuthal anisotropy of particle emission relative to the reaction plane.

In p-p collisions, significant v_2 values have been measured, suggesting the presence of collective flow.

Long-Range Correlations

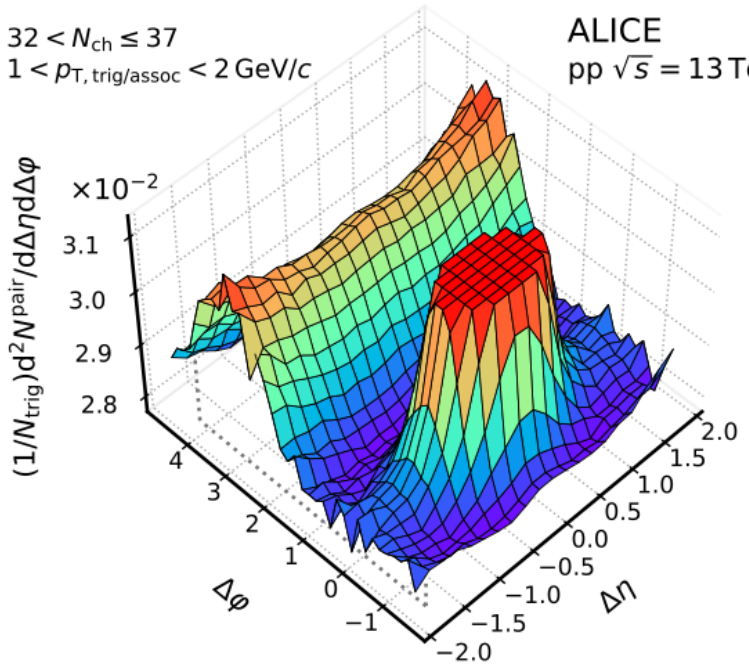

Figure 1.10: Results indicating the formation of nearside ridge signal in high energy p-p collisions [36].

Another indication of collectivity is the presence of long-range rapidity correlations, often referred to as the "ridge" phenomenon. These correlations are seen as a near-side ridge in two-particle correlation functions, extending over a wide range in rapidity [36]. Such structures are typically associated with collective effects in the medium created in the collision. In heavy-ion collisions, such long-range correlations are well-described by hydrodynamic models where the system behaves like a nearly perfect fluid. The pressure gradients in the initial state of the collision lead to collective motion, causing particles to flow together. This collective flow generates correlations that extend over large pseudorapidity ranges, producing the near-side ridge.

Strangeness Enhancement

Strangeness enhancement, which is the increased production of strange hadrons relative to non-strange hadrons, is another signature of QGP formation. This is due to the high mass of strange particles which requires higher energy thresholds for their production. Recent measurements in high-multiplicity p-p collisions have shown an enhancement of strange particle yields, similar to what is observed in heavy ion collisions [37].

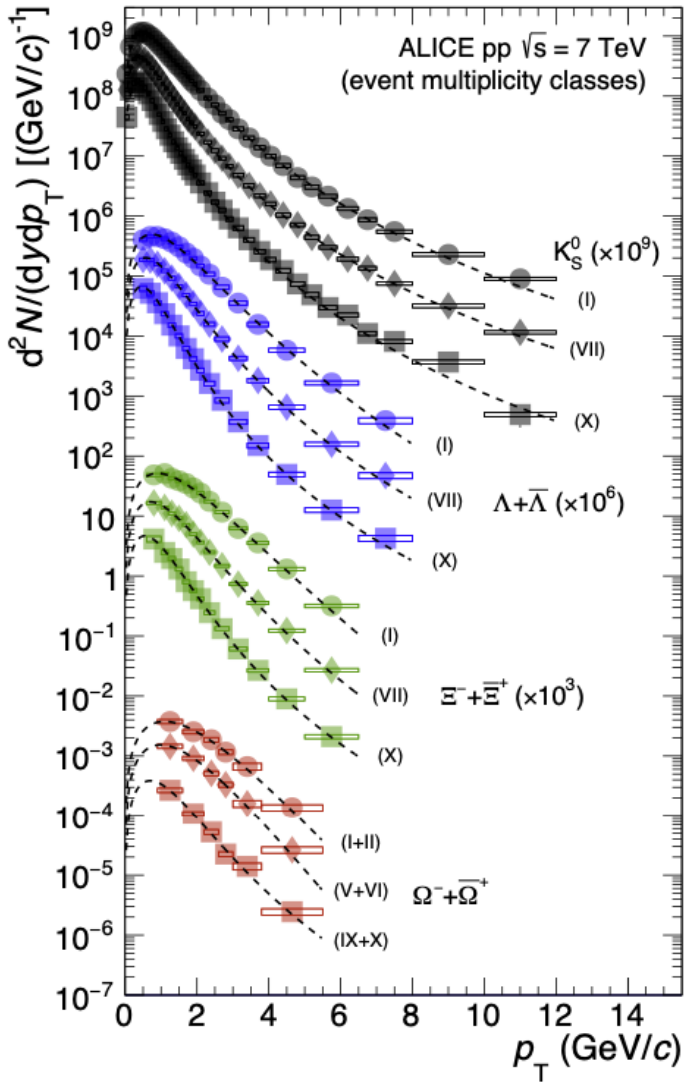

Figure 1.11: Results indicating an enhancement in the production of strange particles for different centrality classes. ([37].

Implications for QGP Formation in Small Systems

The observation of collective behavior and other QGP-like signatures in high-energy p-p collisions challenges the traditional understanding of QGP formation. These results suggest that even small systems can exhibit collective effects and that the formation of a QGP might be more universal than previously thought.

Understanding the conditions under which QGP forms in small systems is an active area of research. It involves studying the initial state effects, the role of gluon saturation, and the mechanisms of thermalization in small volumes. These studies are essential for a complete understanding of QCD matter and the properties of the QGP.

1.6.3 Hadronization

Hadronization is the process by which quarks and gluons transform into hadrons as the QGP cools and undergoes a phase transition. Hadronization is believed to by an analytic cross over at low μ_B and a first order phase transition at high μ_B. This complex process is driven by the confinement property of Quantum Chromodynamics (QCD), which dictates that quarks and gluons cannot exist freely at low temperatures. As the QGP expands and its temperature drops below the critical temperature, quarks and gluons combine to form color-neutral hadrons. This transition preserves certain quantum numbers such as baryon number and strangeness, leading to the production of a variety of hadrons, including pions, kaons, and protons. The detailed dynamics of hadronization are sensitive to the local temperature, density, and flow of the medium, and are crucial for understanding the final particle yields and spectra observed in heavy-ion collision experiments. The statistical hadronization model[8] [38] incorporate these factors to predict particle abundances and provide insights into the properties of the QGP and the nature of the strong force at high temperatures.

[8] A thermodynamic model using the grand-canonical partition function to describe the system.

1.7 Monte-Carlo Modeling (PYTHIA)

Modeling particle collisions is a fundamental aspect of understanding the underlying physics and interpreting experimental results. Monte Carlo modeling is a fundamental computational technique used in particle physics to simulate and analyze particle collisions [39]. It employs random sampling to model the probabilistic nature of physical processes. The basic principle involves generating random numbers to sample from probability distributions governing the interactions of particles. For example, to estimate an integral I over a domain D, one can use the Monte Carlo method:

$$I \approx \frac{1}{N} \sum_{i=1}^{N} f(x_i) \tag{1.33}$$

where N is the number of random samples, x_i are the randomly chosen points in D, and $f(x)$ is the function being integrated. In the context of particle collisions, Monte Carlo methods simulate events by sampling from the differential cross-section $\frac{d\sigma}{d\Omega}$, which describes the likelihood of particles scattering into a specific solid angle Ω. The simulated events provide a statistical representation of physical processes, enabling researchers to predict experimental outcomes and interpret data from particle detectors.

This section focuses on the Monte Carlo based PYTHIA event generator. [9]PYTHIA is a widely used event generator for simulating high-energy collisions involving protons, electrons, and heavy ions. It provides a detailed description of the complex processes involved in particle collisions, from the initial hard scatterings to the final hadronization. PYTHIA incorporates various physical phenomena, including multi-parton interactions (MPI) and color reconnection [40].

[9]It should be noted that current Monte-Carlo event generators do not take into account entanglement. We believe however that specific elements of the used approaches mimic the effects of entanglement allowing them to describe many data well.

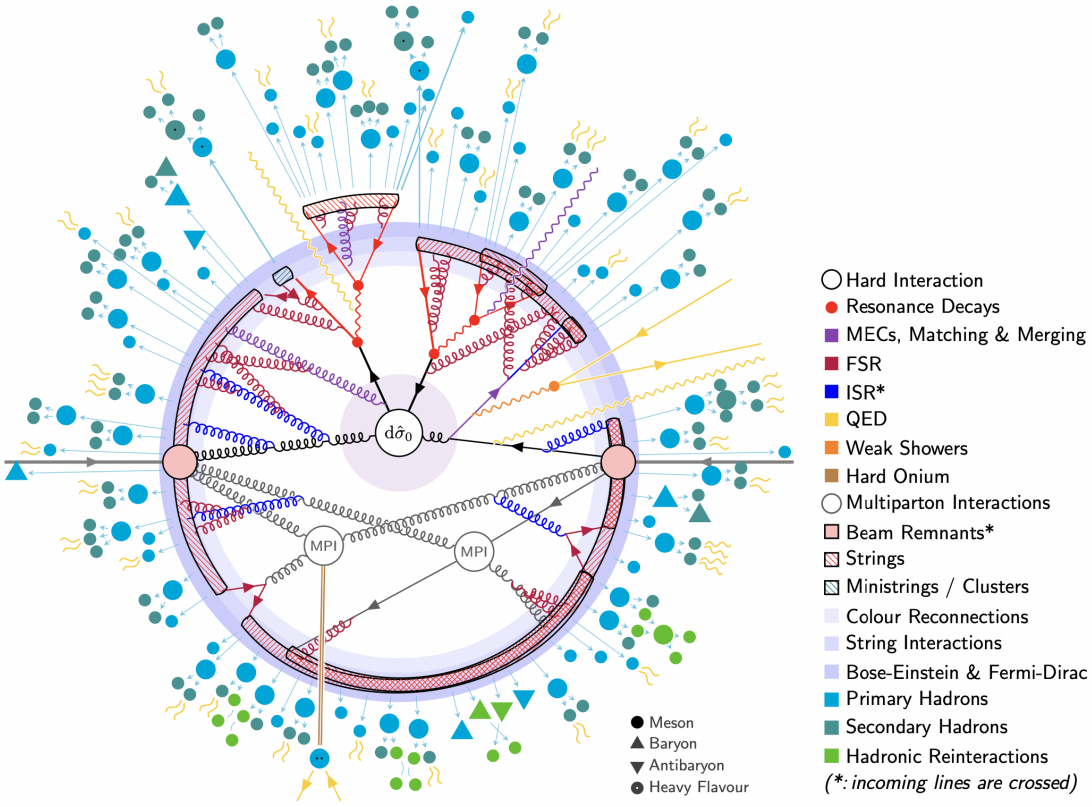

Figure 1.12: Schematic showing a pp event generated in PYTHIA and how particles are produced using different modeled phenomena [40]

1.7.1 Multi-Parton Interactions (MPI)

MPI refers to the theorized occurrence of multiple hard and semi-hard partonic scatterings in a single pp collision. As the collision energy increases, the probability of MPI increases due to the higher parton density at small momentum fractions. MPI contributes to various observables, including the multiplicity of final-state particles, the underlying event activity, and jet production rates.

Modeling MPI in PYTHIA

PYTHIA models MPI by including multiple parton scatterings in the event generation process. The implementation involves several steps:

- **Initial-State Parton Densities**: PYTHIA uses parton distribution functions (PDFs) to describe the momentum distribution of partons within the protons.
- **Parton Scattering**: Multiple parton-parton scatterings are simulated, each contributing to the overall event structure.
- **Impact Parameter**: The spatial overlap of the colliding protons is considered, influencing the number of MPIs.
- **Parton Showers and Hadronization**: Each parton scattering initiates parton showers, which evolve into hadrons through fragmentation and hadronization processes.

1.7.2 Color Reconnection

Color reconnection is a phenomenon that is believed to occur during the hadronization process, where partons from different scatterings can recombine, affecting the final particle correlations and distributions.

Mechanism and Effects

During hadronization, the color charges of partons must neutralize to form color-singlet hadrons. Color reconnection allows partons from different scatterings to rearrange their color connections, leading to the formation of hadrons with different kinematic properties than those produced without reconnection.

Color reconnection can significantly impact observables such as:

- **Jet Structure:** The internal structure of jets can be altered due to the recombination of partons from different sources.

- **Flow-like Patterns:** In high-multiplicity events, color reconnection can lead to flow-like patterns in the final-state particle distributions, resembling collective effects seen in heavy ion collisions.

Modeling Color Reconnection in PYTHIA

PYTHIA includes various models for color reconnection, allowing for the adjustment of parameters to better match experimental data. The implementation involves:

- Reconnection Probability: A probability function determines the likelihood of color reconnection based on the spatial and momentum configurations of partons.

- Reconnection Schemes: Different schemes describe how partons are reconnected to form color-singlet systems. These can include leading order color reconnection which combines pairs of string (color + anti-color), and corrections beyond leading order combining 3 or more strings.

1.8 Hydrodynamic Models

Current hydrodynamic models provides a robust framework for understanding the evolution and particle production in Quark-Gluon Plasma (QGP) created in heavy-ion collisions [41]. These models treat the QGP as a nearly perfect fluid characterized by its energy density, pressure, temperature, and flow velocity fields. The evolution of these fields is governed by the equations of relativistic hydrodynamics, which are derived from the conservation laws of energy, momentum, and baryon number, alongside an equation of state (EoS) that relates pressure to energy density. Initially, the system formed in a heavy-ion collision is highly non-equilibrium; however, it rapidly thermalizes, and hydrodynamic behavior emerges. As the QGP expands and cools, the fluid undergoes collective flow, characterized

by anisotropic expansion due to initial spatial asymmetries. Particle production occurs at the freeze-out stage, where the fluid decouples into individual hadrons. There are two types of freeze-out: chemical freeze-out, where inelastic collisions cease and particle abundances are fixed, and thermal (or kinetic) freeze-out, where elastic collisions stop, and the momentum distributions are set. These models can describe the spectra and flow patterns of the produced particles, offering insights into the transport properties of the QGP, such as viscosity and thermal conductivity, and resolving the complex dynamics of the QGP's evolution from its formation to the final state hadrons observed in detectors.

Hydrodynamic models rely on the system being in thermal equilibrium very early on in the evolution of the collision ($\sim$ 1 fm/c). How this thermalization occurs is unclear in the hydrodynamic picture. Thermalization through entanglement is one proposed solution to this problem.

Calculating the thermodynamic entropy of the particle spectra one can define a temperature of the fireball at the phase transition. Experimental results have shown this temperature to be $T_{CF} = 156.6 + -1.7$ MeV, which agrees almost perfectly with QCD calculations of the chemical freeze-out temperature [42].

1.9 Color Glass Condensate

The Color Glass Condensate (CGC) is a theoretical framework used to describe the high-density regime of gluons in a hadron or nucleus at very high energies. The CGC provides a picture of the initial state of matter in high-energy collisions, where the density of gluons becomes so large that their interactions cannot be neglected. This leads to a state where the gluons form a dense, coherent field that can be described by classical color fields. In a string model where color flux tubes are generated between partons this would be considered flux tube merging.

The CGC framework is based on the separation of scales between the fast-moving valence quarks and the slow-moving gluons. The fast-moving components are treated as static color sources, while the slow-moving gluons are described by classical color fields. This separation allows for a simplification of the problem using effective field theory techniques [43].

1.9.1 Saturation and the Saturation Scale

One of the key concepts in the CGC framework is the idea of saturation. At high energies, the density of gluons increases to the point where non-linear effects become significant. This leads to the saturation of the gluon distribution, where the number of gluons per unit area reaches a maximum limit.

The saturation scale Q_s is a measure of the momentum scale at which saturation occurs. It is defined as the scale where the gluon density per unit transverse area becomes of order $1/\alpha_s$, where α_s is the strong coupling constant. The saturation scale grows with increasing energy, and is inversely proportional to the Bjorken-x.

1.9.2 JIMWLK and BK Equations

The evolution of the CGC with energy is described by the JIMWLK (Jalilian-Marian, Iancu, McLerran, Weigert, Leonidov, and Kovner) and BK (Balitsky-Kovchegov) equations. These are non-linear renormalization group equations that describe the change in the gluon distribution as a function of the energy scale [43].

The JIMWLK equation is a functional differential equation that governs the evolution of the weight functional $W[Y, \rho]$, which describes the probability distribution of the color charge density ρ at rapidity Y [44]:

$$\frac{\partial W[Y,\rho]}{\partial Y} = H_{\text{JIMWLK}} W[Y,\rho], \tag{1.34}$$

where H_{JIMWLK} is the JIMWLK Hamiltonian, which includes terms representing the non-linear interactions between gluons.

The BK equation is a mean-field approximation to the JIMWLK equation and describes the evolution of the dipole scattering amplitude $N(x_{01}, Y)$, where x_{01} is the dipole size [44]:

$$\frac{\partial N(x_{01}, Y)}{\partial Y} = \frac{\alpha_s N_c}{2\pi^2} \int d^2x_2 \frac{x_{01}^2}{x_{02}^2 x_{12}^2} [N(x_{02}, Y) + N(x_{12}, Y) - N(x_{01}, Y) - N(x_{02}, Y)N(x_{12}, Y)], \tag{1.35}$$

where N_c is the number of colors.

1.9.3 Formation of CGC in High-Energy Particle Collisions

The formation of a CGC in high-energy particle collisions, such as those occurring at the Large Hadron Collider (LHC) and Relativistic Heavy Ion Collider (RHIC), involves the following steps:

1. Initial State: Before the collision, the incoming hadrons or nuclei are described by their parton distribution functions (PDFs). At very high energies, the gluon densities in the incoming nuclei become very large, leading to the formation of a dense gluonic state.

2. Gluon Saturation: As the energy increases, the gluon densities reach the saturation scale Q_s. At this point, the gluon distribution becomes non-linear, and saturation effects dominate. The high density of gluons forms a coherent color field, described by the CGC.

3. Classical Color Fields: The CGC can be described by classical color fields due to the high occupancy of gluons. These fields are solutions to the classical Yang-Mills equations with the color sources provided by the fast-moving partons:

$$[D_\mu, F^{\mu\nu}] = J^\nu, \tag{1.36}$$

where D_μ is the covariant derivative, and J^ν is the color current.

4. Collision and Evolution: During the collision, the classical color fields interact and evolve. The evolution is governed by the non-linear dynamics of QCD, and can be described by the JIMWLK and BK equations.

5. Thermalization and Hadronization: As the system evolves, the classical fields eventually thermalize, leading to the formation of a quark-gluon plasma (QGP). The QGP then hadronizes into the final state particles observed in detectors.

1.9.4 Implications and Applications

The CGC framework has important implications for our understanding of high-energy QCD and the initial conditions for the formation of the QGP. It provides a unified description of the initial state of hadronic collisions, linking the parton distributions at low momentum fractions to the observed particle production.

The CGC is also used to predict various observables in high-energy collisions, such as multiplicity distributions, transverse momentum spectra, and azimuthal correlations. These predictions are essential for interpreting experimental data and for understanding the dynamics of high-energy QCD.

In summary, the CGC provides a comprehensive framework for describing the high-density regime of QCD in high-energy collisions. It plays a crucial role in understanding the initial state of matter in such collisions and the subsequent formation and evolution of the QGP. Furthermore, it describes a fully entangled state in which all of the interacting partons become entangled with one another. In the longitudinal direction of the beam this creates a volume in which the only source of entanglement entropy exists on the boundaries of the volume.

1.10 Particle Production Through Color-Dipoles

Kharzeev et. al. established a method for estimating entanglement entropy and mapping the initial-state distribution to the final-state distribution using a 1+1[10] dimensional toy model [45]. This model utilizes the dipole representation, where partons are represented by color dipoles. The evolution of these dipoles is governed by the Balitsky-Kovchegov (BK) equation. In this simplified model, we consider fixed dipole sizes, reducing the BK equation to the BFKL form for the dipole scattering cross-section σ at rapidity Y:

$$\frac{d\sigma(Y)}{dY} = \Delta\sigma(Y), \tag{1.37}$$

where Δ is the BFKL intercept. This equation describes the exponential increase of the cross-section with energy, expressed as $\sigma(Y) \sim \exp(\Delta Y) = (1/x)^{\Delta}$.

To describe the probability $P_n(Y)$ of finding n dipoles at rapidity Y, we use a recurrent equation:

$$\frac{dP_n(Y)}{dY} = -\Delta n P_n(Y) + (n-1)\Delta P_{n-1}(Y). \tag{1.38}$$

This equation captures the depletion of n dipoles due to splitting and the growth from $(n-1)$ dipoles splitting into n dipoles. Introducing the generating function $Z(Y,u)$ helps in solving this cascade:

$$Z(Y,u) = \sum_n P_n(Y) u^n, \tag{1.39}$$

with initial and boundary conditions $Z(Y=0,u) = u$ and $Z(Y,u=1) = 1$. The evolution of $Z(Y,u)$ follows:

[10] One spatial dimension + time

$$\frac{\partial Z(Y,u)}{\partial Y} = -\Delta u(1-u)\frac{\partial Z(Y,u)}{\partial u}. \tag{1.40}$$

The solution to this equation, considering initial and boundary conditions, takes the form:

$$Z(Y,u) = \frac{ue^{-\Delta Y}}{1+u(e^{-\Delta Y}-1)}. \tag{1.41}$$

From this generating function, the probability distribution $P_n(Y)$ can be derived:

$$P_n(Y) = e^{-\Delta Y}(1-e^{-\Delta Y})^{n-1}. \tag{1.42}$$

This model provides a clear framework to understand parton evolution and the corresponding entanglement entropy in high-energy QCD processes.

To calculate the entropy S, we use the von Neumann entropy formula for the probability distribution $P_n(Y)$:

$$S = -\sum_n P_n(Y)\ln P_n(Y). \tag{1.43}$$

Substituting the expression for $P_n(Y)$, and assuming large ΔY the entropy simplifies to:

$$S = \Delta Y. \tag{1.44}$$

The number of produced hadrons, N_{hadrons}, grows exponentially with the rapidity Y as $N_{\text{hadrons}} \sim e^{\Delta Y}$. Therefore, taking the natural logarithm of both sides gives:

$$S \sim \ln(N_{\text{hadrons}}). \tag{1.45}$$

The generating function $Z(Y,u)$ is also related to the multiplicity distributions. By

evaluating the generating function and its derivatives, we can obtain the moments and cumulants of the multiplicity distribution. For example, the average number of partons $\langle n \rangle$ can be derived from the first derivative of $Z(Y, u)$ with respect to u at $u = 1$:

$$\langle n \rangle = u \frac{\partial Z(Y, u)}{\partial u}\bigg|_{u=1} = e^{\Delta Y}. \tag{1.46}$$

Higher-order moments and cumulants, which describe the shape and spread of the multiplicity distribution, can similarly be obtained from higher derivatives of the generating function. This allows the model to relate the statistical properties of particle production directly to the underlying QCD dynamics.

1.11 Thermodynamic-like behavior

It has been discussed how the phases of matter and particle yields in high energy hadronic collisions can be modeled using hydrodynamics. These hydrodynamic models have been successful at describing many phenomena. A new method for understanding the behavior of these interactions has also been introduced using a holographic AdS-CFT model. This section will focus on how thermalization is reached through entanglement, highlighting experimental evidence of a system reaching thermal equilibrium through entanglement and how this process can be quantified. Furthermore, other statistical thermal models have proved effective in prediciting particle production [46].

1.11.1 Thermalization through entanglement

The process of a system reaching a state of thermal equilibrium is called thermalization. Hydrodynamic models used to understand QGP behavior require the system to be in thermal equilibrium. We do not fully understand how this thermalization occurs. The

current hydrodynamic description requires thermalization to be reached in a time less than 1 fm/c [35]. Typically, in normal phases of matter like a gas of particles thermal equilibrium is reached through particle interactions. However, these reactions take some time, if interactions caused thermalization in such a fast time causality would be broken. Therefore because the fireball thermalizes so early it is unlikely that thermal equilibrium is reached through particle interactions alone.

We have now established that entanglement exists in the pure state of a proton. We have also established that if we provide enough energy to a proton such that a state of deconfinement is reached a quark-gluon plasma will form. Furthermore, we understand that this QGP should thermalize early in its evolution. Could the entanglement of this initial state play a role in the thermalization of the QGP? To answer such a question we must first establish a connection between entanglement and thermalization.

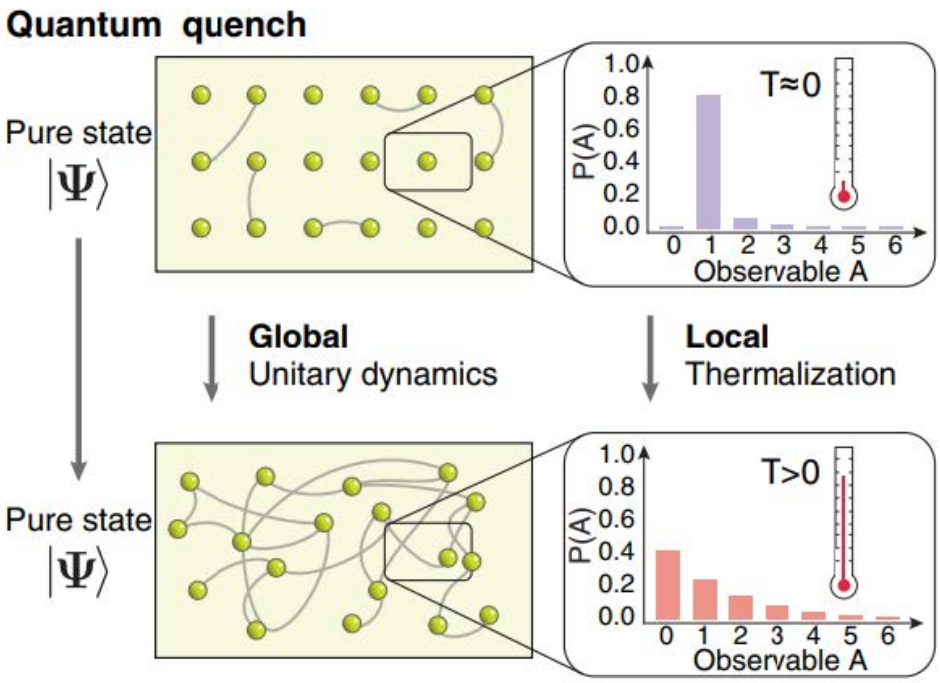

Figure 1.13: A lattice of entangled rubidium-87 atoms reaches instant thermalization through entanglement [47]

A recent study done at Harvard University has demonstrated that it is possible for a system, devoid of particle interactions, to reach thermal equilibrium through entanglement [47]. In their experiment a group of entangled rubidium-87 is placed in a two dimensional lattice. Initially this state is set to a temperature very near 0 K. Then a local temperature increase, or quench, is introduced to one element of the system. Following the quench atoms redistributed themselves within the lattice through tunneling. This new distribution of particles behaved like a thermal distribution and the system was in a global state of thermal equilibrium. Because there were no particle interactions decoherence did not occur and they system remained a pure quantum state. Thus thermalization occurred through entanglement which is an instantaneous phenomena allowing for instantaneous thermalization.

Entanglement in trapped ion experiments such as this is highly controlled and quantified. This is achievable through high precision control over EM fields and maintaining an extremely low temperature so that there is no loss of coherence to the environment. This is not reflective of the high energy density environment of the QGP. Next, we will look at an experimental result demonstrating a retention of coherence in a "hot" and "dense" medium interacting strongly.

1.11.2 Entanglement in a hot dense medium

In a recent study published in Nature, researchers investigated the generation and preservation of entanglement in a hot alkali vapor, made up of rubidium-87 (^{87}Rb) atoms, under conditions dominated by random spin-exchange collisions [48]. Their approach utilizes optical quantum non-demolition (QND) measurement techniques to produce entangled states in a strongly interacting atomic medium, demonstrating that high temperatures and strong random interactions do not necessarily destroy many-body quantum coherence.

The experimental setup, as shown in Figure 1.14(a), consists of a glass cell containing hot

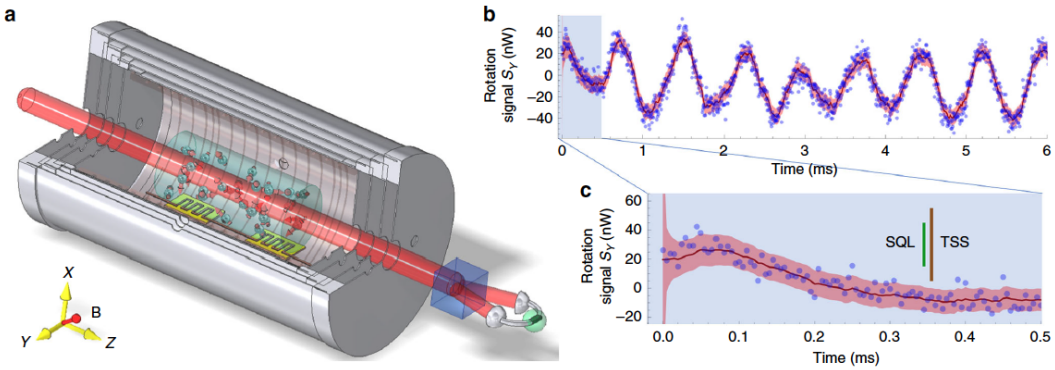

Figure 1.14: Experimental setup and measurement of entanglement in a hot dense medium [48]

^{87}Rb vapor along with 100 torr of nitrogen buffer gas. A linearly polarized probe beam, red-detuned by 44 GHz from the ^{87}Rb D1 line, passes through the vapor cell. The transmitted light is detected using a shot-noise-limited polarimeter, which measures the projection of the collective spin on the probe direction, plus optical shot noise. A static magnetic field is applied along the [1, 1, 1] direction, causing the spin components to precess.

We perform continuous, non-destructive readout of the spin polarization using Faraday rotation of the off-resonance light. The observed polarization rotation gives insight into the spin dynamics, which are described by the Langevin equation:

$$d\mathbf{F} = \gamma \mathbf{F} \times \mathbf{B}\, dt - \Gamma \mathbf{F}\, dt + \sqrt{2\Gamma Q}\, d\mathbf{W}, \tag{1.47}$$

where $\mathbf{F}$ is the collective spin, γ is the gyromagnetic ratio, $\mathbf{B}$ is the magnetic field, Γ is the relaxation rate, Q is the equilibrium variance, and $d\mathbf{W}$ represents the Wiener increments.

Figure (b) illustrates the rotation signal $S_y(t)$, showing clear atomic spin coherence over millisecond time scales. The spin squeezing parameter ξ^2 is used to detect entanglement, defined as:

$$\xi^2 = \frac{\Delta F^2}{\mathrm{SQL}}, \tag{1.48}$$

where ΔF^2 is the total spin variance and SQL is the standard quantum limit. A value of $\xi^2 < 1$ indicates the presence of entanglement.

Results demonstrate that at least $1.52(4) \times 10^{13}$ of the $5.32(12) \times 10^{13}$ participating atoms are entangled, forming a macroscopic singlet state that persists far beyond the spin-thermalization time of the vapor. Figure (c) shows the early signal where the Kalman filter rapidly acquires a sub-SQL estimate for F_z within 20 microseconds, much less than the coherence time.

These findings indicate that QND measurement can generate and preserve entanglement in hot atomic systems, even in the presence of strong local interactions. The hot, dense atomic medium investigated in the study shares similarities with a quark-gluon plasma (QGP) in that both systems exhibit strong interactions among their constituent particles and can maintain collective quantum behaviors under extreme conditions. In both the hot atomic vapor and QGP, particles interact frequently and strongly, leading to complex many-body dynamics that preserve entanglement or coherence over significant distances and timescales, despite the high level of thermal agitation. These characteristics make the study of such systems crucial for understanding the fundamental properties of matter under extreme conditions.

1.11.3 Thermodynamic Entropy

Thermodynamic entropy is a fundamental concept in statistical mechanics and thermodynamics, quantifying the degree of disorder or randomness in a system. It is a measure of the number of microscopic configurations that correspond to a macroscopic state.

Thermodynamic entropy S is closely related to temperature T through the fundamental equations of thermodynamics. Entropy quantifies the amount of disorder or randomness in a system, and its relationship with temperature is central to understanding heat exchange

and energy distribution. The change in entropy dS with respect to an infinitesimal amount of heat dQ added to the system at temperature T is given by [49]:

$$dS = \frac{dQ}{T}. \tag{1.49}$$

For a reversible process, the total change in entropy can be expressed as:

$$\Delta S = \int \frac{dQ_{\text{rev}}}{T}, \tag{1.50}$$

where dQ_{rev} denotes the heat added in a reversible process. This relationship indicates that entropy increases when heat is added to the system, and this increase is inversely proportional to the temperature. Higher temperatures lead to smaller changes in entropy for the same amount of heat added, reflecting that systems at higher temperatures have higher energy and thus smaller relative increases in disorder.

Moreover, the second law of thermodynamics states that for any spontaneous process, the total entropy of the system and its surroundings always increases. This can be written as:

$$\Delta S_{\text{total}} = \Delta S_{\text{system}} + \Delta S_{\text{surroundings}} \geq 0. \tag{1.51}$$

These fundamental equations illustrate the intrinsic connection between entropy and temperature, underlying the principles of heat transfer, energy distribution, and the direction of spontaneous processes in thermodynamics.

Mathematical Definition of Thermodynamic Entropy

The thermodynamic entropy S of a system can be defined using the Boltzmann entropy formula for a probability distribution of states. If a system can exist in different microstates

i with probabilities p_i, the entropy is given by [49]:

$$S = -k_B \sum_i p_i \ln p_i, \tag{1.52}$$

where k_B is the Boltzmann constant, and the sum is taken over all possible microstates of the system.

This equation can be derived from the Boltzmann entropy formula:

$$S = k_B \ln \Omega, \tag{1.53}$$

where Ω is the number of accessible microstates. For a system in which each state has an equal probability $p_i = \frac{1}{\Omega}$, this reduces to the aforementioned definition.

Properties of Thermodynamic Entropy

Thermodynamic entropy has several important properties:

- **Additivity:** The entropy of a composite system is the sum of the entropies of its subsystems.
- **Extensivity:** Entropy is an extensive property, meaning it scales with the size of the system.
- **Second Law of Thermodynamics:** In any spontaneous process, the total entropy of an isolated system always increases or remains constant.

1.11.4 Relating Thermal Entropy to Entanglement Entropy

Thermodynamic entropy measures the disorder in a macroscopic system based on the distribution of microstates. It is proportional to the Shannon entropy which is defined

as [50] [11]:

$$S = -\sum_i p_i \ln p_i, \tag{1.54}$$

It reflects the number of ways a system can be arranged microscopically while appearing the same macroscopically. The higher the number of accessible microstates, the higher the entropy.

Entanglement entropy, on the other hand, measures the degree of quantum entanglement between subsystems. For a bipartite quantum system described by a pure state ψ, the entanglement entropy of subsystem A is given by the von Neumann entropy of the reduced density matrix ρ_A [51]:

$$S_A = -\mathrm{Tr}(\rho_A \log \rho_A). \tag{1.55}$$

Here, entropy is obtained by tracing out the degrees of freedom of subsystem B. Each definition of entropy is conceptually and mathematically similar though they are defined under a different context.

Despite their differences, both entropies quantify disorder within a system. Thermodynamic entropy measures disorder in terms of the probability distribution of microstates, while entanglement entropy measures disorder in terms of quantum correlations between subsystems.

In a thermodynamic context, entropy increases as the system approaches equilibrium, representing the spread of energy and matter into available states. In a quantum context, entanglement entropy increases as subsystems become more entangled, representing the spread of quantum information across the system.

Both concepts highlight the fundamental nature of entropy as a measure of uncertainty

[11] This is the definition of entropy we will use to describe the final state of produced hadrons

or lack of information about the exact state of a system. They provide complementary perspectives on the nature of disorder and information in physical systems, linking macroscopic thermodynamic properties to microscopic quantum behaviors.

1.12 Quantum Information

Quantum information is a revolutionary field at the intersection of quantum mechanics and information theory, exploring how quantum systems can be used to represent, store, and manipulate information. Unlike classical information, which is encoded in binary bits (0s and 1s), quantum information is encoded in quantum bits, or qubits. These qubits leverage the principles of superposition and entanglement, allowing them to exist in multiple states simultaneously and to be intricately correlated with each other, respectively. This enables quantum computers to perform certain computations exponentially faster than classical computers and promises profound advancements in cryptography, communication, and computation. In our context we are using the formalism of quantum information to describe the initial conditions and evolution of the partonic system. Initially the information is encoded in the distribution of partons and transferred to a distribution of produced hadrons. We define the information in the particle basis however there exist other basis of information that contribute the overall entanglement entropy that are difficult to define at an experimental level. Because we lack complete information about the system we cannot measure a true entanglement entropy but rather a portion of the entanglement entropy known in literature as the "entropy of ignorance" [33, 52].

1.12.1 Shannon Entropy

Shannon entropy is a fundamental concept in information theory that quantifies the uncertainty or randomness in a set of possible outcomes. It was introduced by Claude Shannon in his seminal 1948 paper "A Mathematical Theory of Communication" and has since become a cornerstone in the fields of data compression, communication theory, and many other areas of science and engineering [50].

Definition and Mathematical Formulation

Shannon entropy H is a measure of the average amount of information produced by a stochastic source of data. For a discrete random variable X with possible outcomes $\{x_1, x_2, \ldots, x_n\}$ occurring with probabilities $\{p_1, p_2, \ldots, p_n\}$, the Shannon entropy is defined as:

$$H(X) = -\sum_{i=1}^{n} p_i \ln p_i, \tag{1.56}$$

where $p_i = P(X = x_i)$ is the probability of the outcome x_i, and the logarithm is taken base e to measure entropy in natural units reflective of a thermodynamic entropy.

Properties of Shannon Entropy

Shannon entropy has several important properties that make it a useful measure of uncertainty:

- **Non-Negativity:** Entropy is always non-negative[12]:

$$H(X) \geq 0. \tag{1.57}$$

- **Additivity:** For independent random variables X and Y, the entropy of their joint

[12]The entropy is zero if and only if the random variable X takes a single value with probability 1.

distribution is the sum of their individual entropies:

$$H(X,Y) = H(X) + H(Y). \tag{1.58}$$

- **Maximum Entropy:** For a given number of outcomes, entropy is maximized when all outcomes are equally likely. For a random variable X with n equally likely outcomes ($p_i = \frac{1}{n}$ for all i), the entropy is:

$$H(X) = \ln n. \tag{1.59}$$

1.12.2 Entropy of ignorance

The entropy of ignorance, arises due to the limited information accessible from experimental measurements. Typically, experimental observables in particle collisions are related to the average number of particles or multi-parton momentum distributions, such as $\langle a^\dagger(k_1)a(k_1)\cdots a^\dagger(k_n)a(k_n)\rangle$. These observables do not provide information about the off-diagonal elements of the density matrix in the number operator basis.

To formalize this, we define the density matrix of ignorance ρ_I by setting all off-diagonal elements of the reduced density matrix to zero:

$$\rho_I = \mathrm{diag}(\rho_A). \tag{1.60}$$

The entropy of ignorance S_I is then given by the von Neumann entropy of this diagonalized density matrix:

$$S_I = -\mathrm{Tr}\left(\rho_I \ln \rho_I\right). \tag{1.61}$$

It has been established that the entropy of ignorance is always greater than or equal

to the entanglement entropy, i.e., $S_I \geq S_A$. This can be understood as a manifestation of the loss of information when the off-diagonal elements of the density matrix, which encode quantum correlations, are ignored. [13]

1.13 Entanglement shown in DIS Experiments

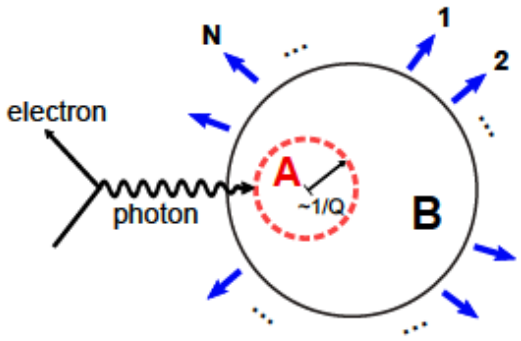

Figure 1.15: In DIS interactions a volume of the proton is probed by the vitual photon. Entanglement entropy arises between region A and region B due to decoherence [51]

Recent analysis have shown some preliminary evidence of entanglement in particle collisions. These studies have shown the correspondence in DIS experiments. To show this correspondence a calculation of entanglement entropy is made between a region of the proton probed by the virtual photon and the remainder of the proton. This calculation is then compared to a Shannon entropy of produced hadrons experimentally measured. One such analysis using data from the H1 DIS experiment showed strong agreement between the initial and final state entropy [53]. This result makes a strong case for entanglement being the driving mechanism behind particle production. However this agreement has yet to be

[13] In our analysis, which relies on experimental measurement, a correction will have to be made to account for these off diagonal elements of the density matrix in the number basis. This correction will be described in a later section.

shown in hadron-hadron collisions where the dynamics and phases of matter are far more complicated.

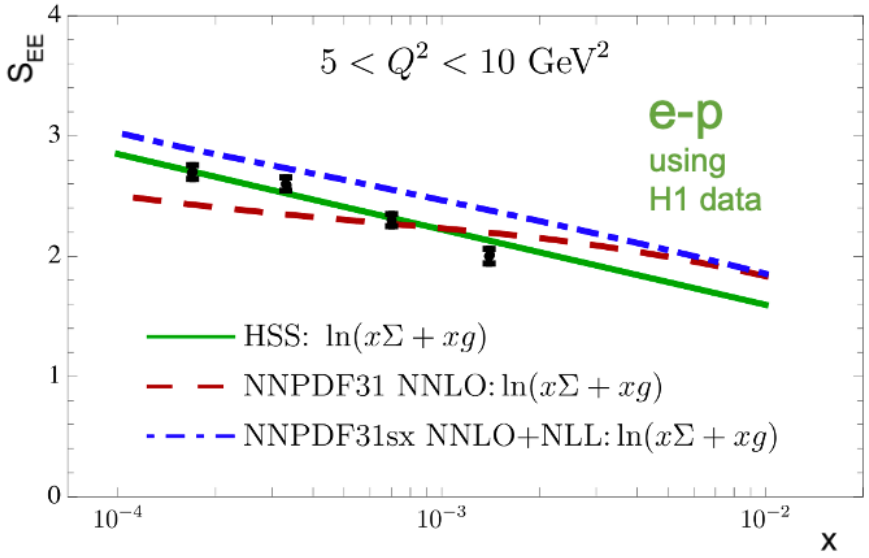

Figure 1.16: Initial-state entropy (lines) compared to final-state entropy (points) as a function of x using H1 data. [53]

1.13.1 Challenges of Measuring Entanglement in more Complex Systems

While strong evidence of entanglement has been shown in DIS experiments showing this in a more complex system like p-p collisions poses some unique challenges. In 2019 a preliminary study was published which helped to establish some of the methods necessary for measuring this in p-p collisions [51].

One challenge in dealing with p-p collisions involves the convolution of two partonic distributions rather than one as defined by initial state PDFs. In order to account for this some assumptions will have to be made about the production of particles and a transformation on the final state will be applied.

Other challenges with measuring entanglement in p-p collisions involve resolution of the kinematic variables Q and x which are more easily obtained in DIS measurements.

An advantage to DIS experiments is the ability to readily calculate and bin data in terms of the energy scale Q (see section on Deep Inelastic Scattering). Since precise information about this scale is needed in order to define the initial state this is a major advantage. In contrast in p-p collisions Q is calculated from theoretical predictions based on BK equations (as described in a previous section).

Additionally the energy scale in p-p collisions is defined by the saturation scale which is estimated to be near one at LHC energies (in the above figure data is binned in scales significantly higher). At low-Q the initial state calculations are less defined and uncertainties in the PDFs become very large.[14]

Finally, at higher values of Q the off diagonal elements of the density matrix are reduced and become negligible [15]. Therefore, the entropy of ignorance and entanglement entropy become approximately equal and no correction needs to be made to account for these unmeasured degrees of freedom.

[14]The difficulty in extrapolating PDFs to low energy scales will be discussed in a later chapter
[15]this will also be shown in chapter 4

Chapter 2

Experimental setup

The distribution in the final state will be measured using the ALICE detector at the LHC. Since we are calculating this in the particle number basis we will use the detector as a counting experiment, quantifying the number of hadronic states produced in high energy collisions. In order to make a meaningful comparison to the kinematics of the initial state we will also use the detector to measure a geometric angle of production called psuedorapidity.[1]

2.1 LHC

The Large Hadron Collider (LHC) is the world's largest and most powerful particle accelerator, located near Geneva, Switzerland. Operated by the European Organization for Nuclear Research (CERN), the LHC is a marvel of modern engineering and scientific collaboration, consisting of a 27-kilometer-long circular tunnel situated approximately 100 meters underground. The LHC is designed to explore the fundamental structure of matter and the forces that govern the interactions between particles, thereby enhancing our understanding of the universe and its underlying principles.

[1] How the pseudorapidity angle relates to the kinematics of the initial state will be discussed in the next chapter

Figure 2.1: The Large Hadron Collider in Geneva, Switzerland

Construction of the LHC began in the late 1990s and was completed in 2008. Since then, the accelerator has been instrumental in a wide range of groundbreaking experiments in particle physics, leading to significant advances in our knowledge of the fundamental forces and constituents of the universe. The LHC works by accelerating beams of protons and other particles to near the speed of light. These beams are guided by powerful superconducting magnets, which operate at extremely low temperatures only a few degrees above absolute zero. As the particles reach their maximum velocities, they are directed into collision courses, resulting in high-energy impacts that probe the particles internal structure and allow physicists to study the fundamental building blocks of matter and the forces that bind them together.

The LHC's impressive capabilities have led to several landmark discoveries in the field of particle physics. One of its most significant achievements was the discovery of the Higgs boson in 2012 [15], a previously elusive particle that is responsible for giving other particles mass through its interaction with the Higgs field. This breakthrough confirmed the existence

of the Higgs mechanism, a key component of the standard model of particle physics, and led to the awarding of the 2013 Nobel Prize in Physics to theoretical physicists Peter Higgs and François Englert.

Apart from the Higgs boson, the LHC has been instrumental in the investigation of exotic particles and rare decay processes, as well as the study of matter-antimatter asymmetry and the search for dark matter candidates. The LHC's high-energy collisions recreate conditions similar to those that existed shortly after the Big Bang, enabling researchers to probe the early universe's evolution and the mechanisms that govern its behavior.

The LHC hosts four major experiments, each focusing on different aspects of particle physics: ATLAS (A Toroidal LHC ApparatuS) [54], CMS (Compact Muon Solenoid) [55], LHCb (Large Hadron Collider beauty) [56], and ALICE (A Large Ion Collider Experiment) [57]. Each of these detectors is designed to study specific aspects of high-energy particle collisions, contributing to a comprehensive understanding of the universe's fundamental constituents and the interactions between them.

ALICE, in particular, is dedicated to investigating the properties of strongly interacting matter under extreme conditions, such as those found in the early universe. This is achieved by examining the behavior of quark-gluon plasma.

2.2 A Large Ion Collider Experiment (ALICE)

ALICE is one of the four major experiments at the Large Hadron Collider (LHC) at CERN. The interaction point assigned to the ALICE detector is known as Interaction Point 2 (IP2) located at 6 o'clock on the LHC's main ring. The ALICE detector is located near the village of St. Genis-Pouilly in France, positioned 56 meters underground. The ALICE collaboration consists of nearly 2000 scientists from 174 physics institutes across 40 countries [58].

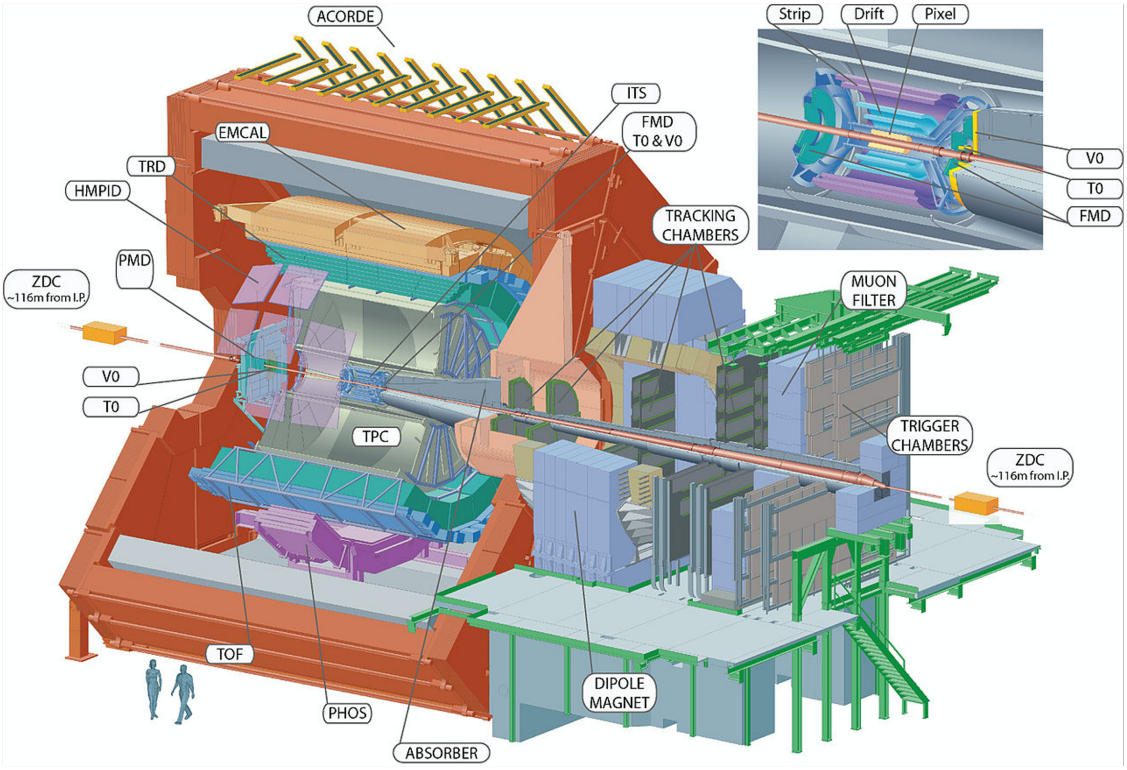

Figure 2.2: A Large Ion Collider Experiment (ALICE) detector schematic

It is specifically designed to study the properties of strongly interacting matter under extreme conditions, such as those found in the early universe. ALICE also plays a crucial role in the study of proton-proton collisions, providing valuable insights into the strong force and the behavior of quarks and gluons that make up protons and neutrons.

The ALICE detector is a complex system comprising various subsystems designed to detect, identify, and measure particles produced in high-energy collisions. These subsystems work together to provide a comprehensive understanding of the collision events and their products. In the context of this work, the primary particle multiplicity, which is a key observable for characterizing the event and measuring the thermodynamic entropy, can be obtained through several subsystems of the detector related to particle tracking and vertex positioning.

2.2.1 Fast Interaction Trigger (FIT)

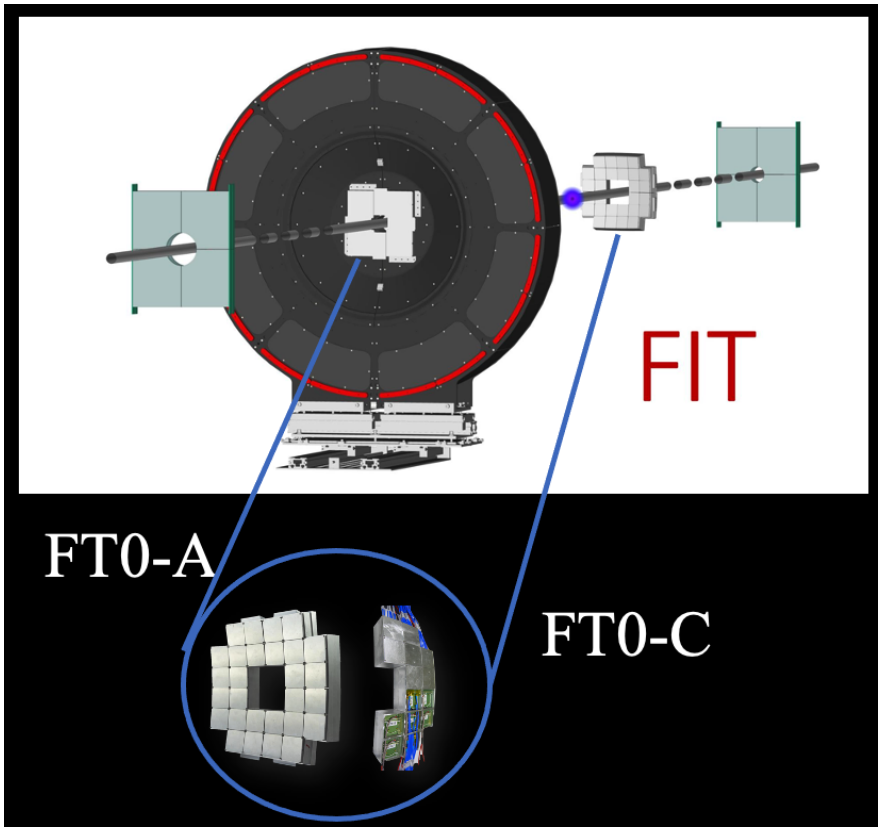

Figure 2.3: Schematic of the FIT detector [59]

The Fast Interaction Trigger (FIT) consists of two arrays of Cherenkov modules, designated FT0-A and FT0-C, positioned approximately 0.8 meters and 3 meters from the interaction point (IP), respectively. These arrays are equipped with 24 and 28 Cherenkov modules and are used for multiple purposes including triggering, event classification, and multiplicity measurements [59].

The FIT covers a pseudorapidity range of $3.8 < \eta < 5.0$ and $-3.4 < \eta < -2.3$, providing time resolution on the order of $\sim 33\,\text{ps}$. This precise timing capability allows the FIT to ensure that collisions are close enough in time to filled bunch crossings, facilitating minimum bias triggers. The Cherenkov modules detect the light produced when charged particles pass through the detector medium, allowing for fast and efficient identification of collision events.

The FIT is essential for synchronizing the data acquisition with the actual collision events, ensuring high-quality data for further analysis. The design and functionality of the FIT, as illustrated in the figure, highlight its critical role in the ALICE experiment at the LHC. In this analysis this detector will be primarily used for triggering, ensuring we are looking at high quality events.

2.2.2 Internal Tracking System (ITS)

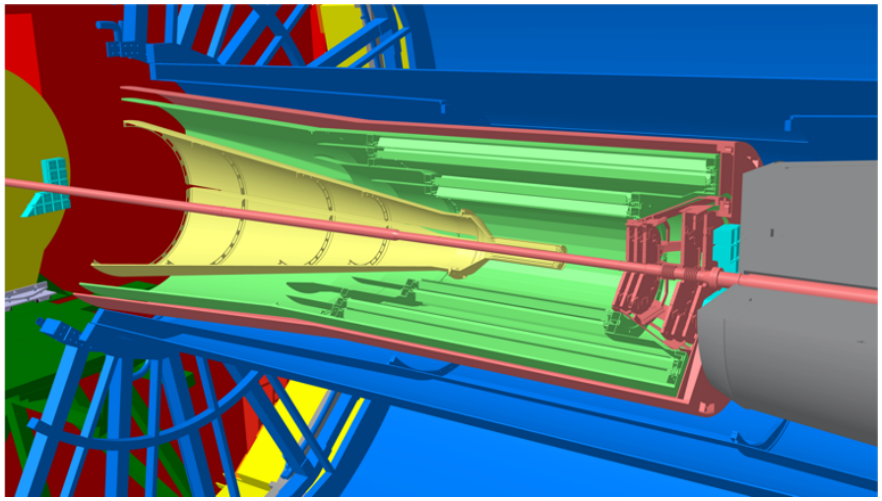

Figure 2.4: Schematic of the ITS featuring seven layers of silicon pixel detectors.

The ITS is another crucial subsystem for particle tracking and vertex positioning in ALICE. It produces short "tracklets" near the interaction point, which greatly improves the accuracy of particle counting. The ITS is a high-resolution, silicon-based detector situated close to the beamline and consists of seven concentric cylindrical layers of highly granulated silicon pixel detectors. This system provides impressive impact parameter resolution, readout rate, and tracking efficiency, especially for low-p_T particles [60].

This subsystem serves several purposes in ALICE, including the precise determination of primary and secondary vertices, tracking of charged particles, and the identification of heavy-flavor hadrons through their decay vertices. With pseudorapidity coverage of $|\eta| < 1.22$, the

ITS is an important subsystem in particle tracking near the interaction point. The ITS features 12.5 billion Monolithic Active Pixel Sensors (MAPS) with a spatial resolution of 5 μm, and operates at a readout rate of 500 kHz to 1 MHz for proton-proton collisions. The high spatial resolution of the ITS allows precise tracking of primary particles and serves as one of the primary subsystems for particle counting in this analysis.

2.2.3 TPC

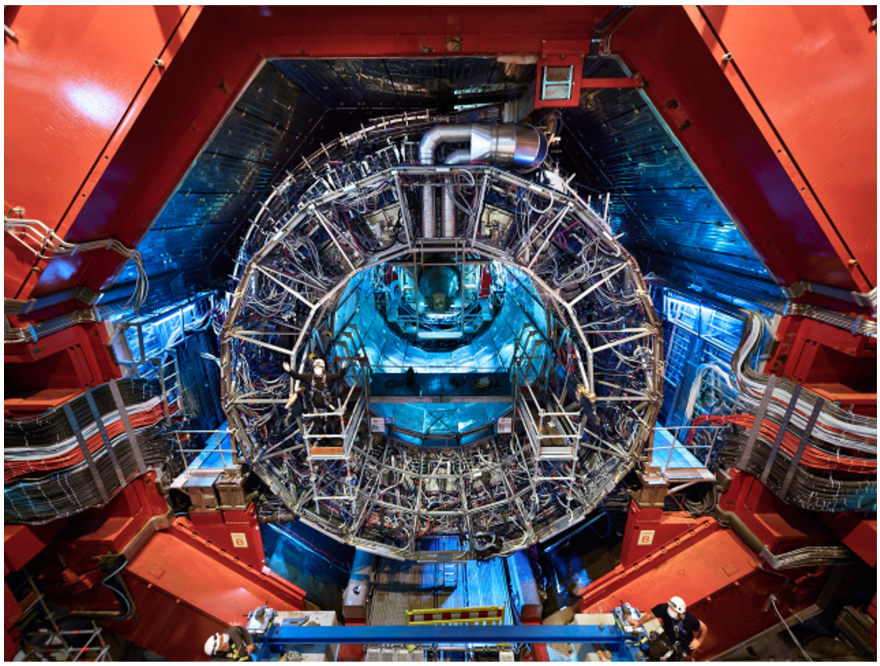

Figure 2.5: Photo of the ALICE detector featuring the central tracking system known as the TPC.

The TPC is the central tracking system in ALICE and plays a pivotal role in measuring charged-particle momenta and tracking primary and secondary vertices. It produces tracks which, along with tracklets from the ITS, provide an accurate count of primary charged particles. The TPC is a large cylindrical drift chamber filled with a gas mixture, typically Neon-based. When charged particles traverse the TPC, they ionize the gas, creating electrons that drift towards the endcaps of the chamber under the influence of an electric field. For

the Run 3 upgrade, the Multi-Wire Proportional Chambers (MWPC) were replaced with Gas Electron Multipliers (GEM), thus removing rate restrictions, reducing ion backflow, and minimizing space-charge distortions [61].

The TPC provides excellent tracking and momentum resolution, as well as good particle identification capabilities through the measurement of specific energy loss (dE/dx) of particles as they traverse the detector. By correlating the energy loss information with the particle's momentum, it is possible to distinguish between different particle species, which is essential for determining the primary particle multiplicity in collision events. The upgraded TPC with GEM technology ensures enhanced performance, allowing for more precise and reliable tracking in the central barrel of ALICE.

2.3 Particle Tracking in ALICE

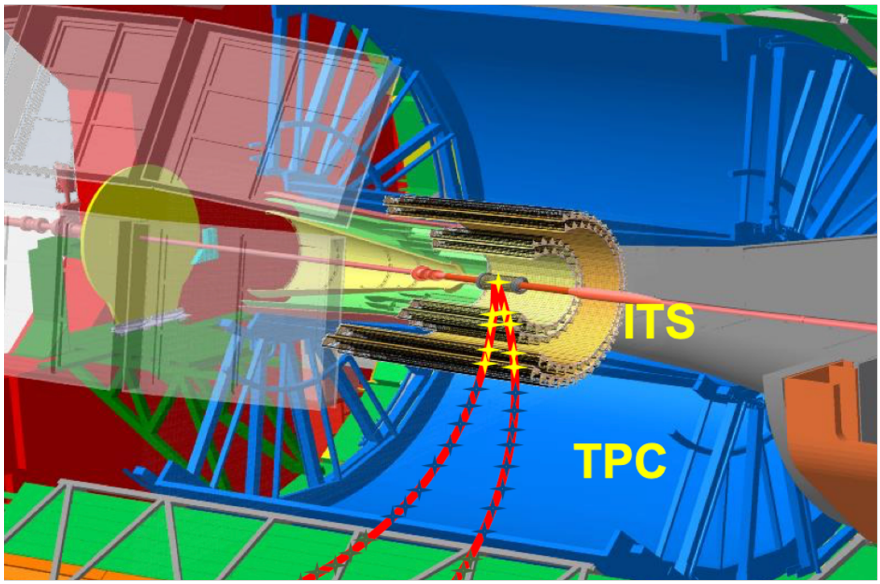

Figure 2.6: Schematic showing how particles are tracked using the two main subsystems for particle counting, the ITS and TPC.

Particle counting is done using the number of good-quality global tracks measured by

the TPC, supplemented by the number of ITS tracklets. The global tracks contribute up to $|\eta| \leq 0.9$, ITS-standalone tracklets cover up to $|\eta| \leq 1.3$. Matching of these tracks and tracklets at mid-rapidity provides the most reliable approach to counting charged particle tracks and defining their angular position [62].

2.4 Quality Control

Data collecting during LHC runs is a complex process requiring consistent monitoring and parameter adjusting of detector subsystems and data acquisition systems. Often failures in the system result in unreliable data. Some of this data is salvageable in that corrections can be made to it, while other data needs to be removed entirely. Data quality assessment is a key aspect of this process and needs to be done both during data taking (synchronous QC) and after data acquisition (asynchronous QC). As a member of the ALICE collaboration I participated in both synchronous and asynchronous QC.

During data taking a series of reports is continuously generated and monitored in real time for data quality issues. While working in ALICE I worked shifts as a QC expert monitoring these reports and reporting issues.

As service work to the collaboration I also worked on software used for generating asynchronous QC reports for the TPC. While working in this capacity I developed and implemented a new method for determining data quality which is now in use by the TPC-QC task force. This method uses the formalism of information theory and entropy to determine data quality and is described in the next section.

2.5 KL-Divergence for Data Quality Assessment

In the context of the Time Projection Chamber (TPC) Quality Control (QC) group, ensuring high data quality is crucial for accurate physics analysis. A method was developed and implemented to assess data quality by comparing a set of histograms. This method leverages the Kullback-Leibler (KL) divergence, a statistical measure used to quantify the difference in information content between two probability distributions [63]. This section highlights the work done to utilize KL divergence for determining data quality by comparing individual histograms to an average histogram.

2.5.1 Theoretical Framework of KL-Divergence

The Kullback-Leibler (KL) divergence, also known as relative entropy, is a measure of how one probability distribution diverges from a second, expected probability distribution. Mathematically, the KL divergence from distribution Q to distribution P is defined as [63]:

$$D_{\mathrm{KL}}(P \parallel Q) = \sum_i P(i) \log \frac{P(i)}{Q(i)}, \tag{2.1}$$

where $P(i)$ and $Q(i)$ are the probability distributions over the same discrete variable i.

In the continuous case, the KL divergence is given by:

$$D_{\mathrm{KL}}(P \parallel Q) = \int_{-\infty}^{\infty} p(x) \log \frac{p(x)}{q(x)}\, dx, \tag{2.2}$$

where $p(x)$ and $q(x)$ are the probability density functions of the continuous random variable x.

2.5.2 Method for Data Quality Assessment

1. **Normalization:** Before calculating the KL divergence, it is essential to normalize the histograms so that they can be interpreted as probability distributions. Each bin in a histogram is normalized by dividing by the total number of entries in the histogram:

$$P_i = \frac{H_i}{\sum_j H_j}, \tag{2.3}$$

where H_i is the bin content of the i-th bin in the histogram and P_i is the normalized bin content.

2. **Histogram Collection:** Collect a set of histograms from different runs of data. Each histogram represents a distribution of some measured quantity in the TPC.

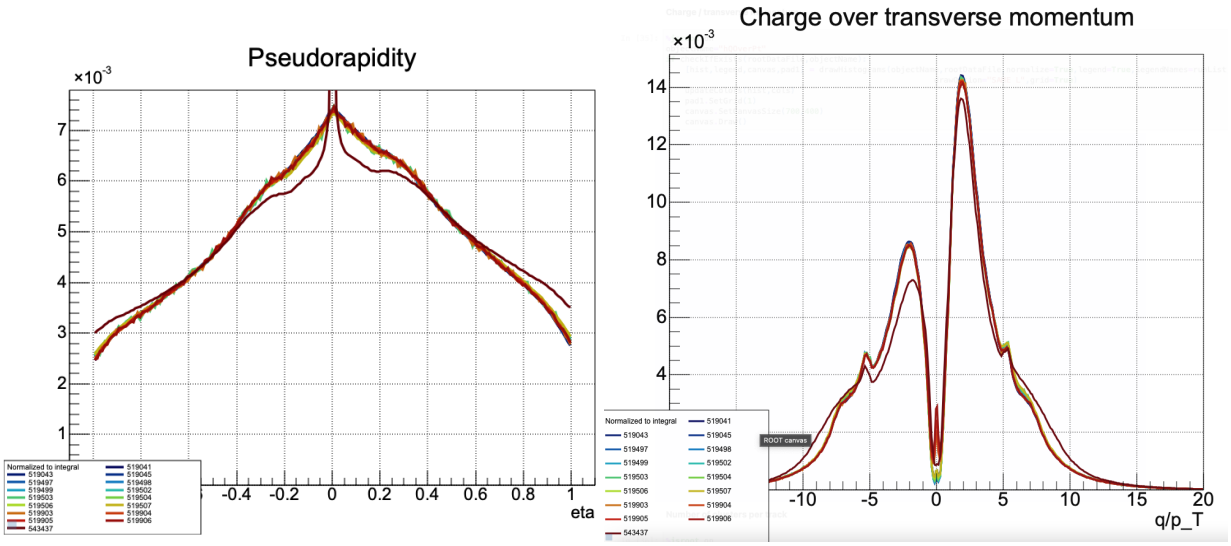

Figure 2.7: a) Example set of data collected from TPC data representing η distributions from separate data taking periods or runs. (b) Example set of data collected from TPC data representing Charge/Transverse momentum distributions from separate data taking periods or runs.

3. **Average Histogram Creation:** Create an average histogram from the set of collected histograms. This average histogram represents the expected distribution of the measured quantity and is defined bin by bin as:

$$\bar{H}(i) = \frac{1}{N} \sum_{k=1}^{N} H_k(i), \tag{2.4}$$

where H_i is the bin content of the i-th bin in the histogram and N represents the total number of histograms in the set.

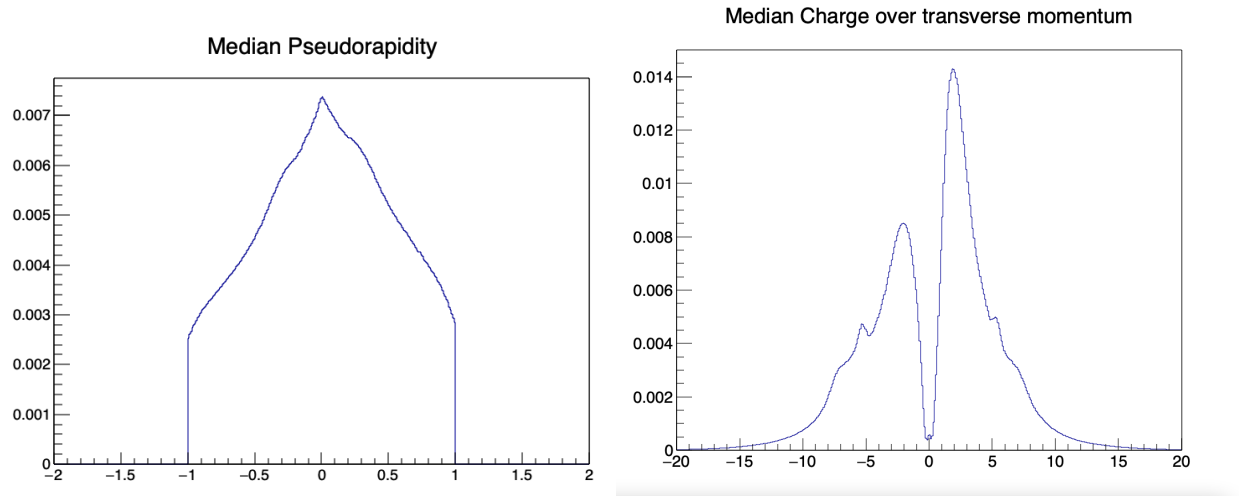

Figure 2.8: Average distributions of η and Charge/Transverse momentum from the entire data set.

4. **KL Divergence Calculation:** Calculate the KL divergence between the average histogram and each individual histogram. This step quantifies the difference between the individual runs and the expected distribution:

$$D_{\mathrm{KL}}(P_{\mathrm{avg}} \parallel P_k) = \sum_i P_{\mathrm{avg},i} \log \frac{P_{\mathrm{avg},i}}{P_{k,i}}. \tag{2.5}$$

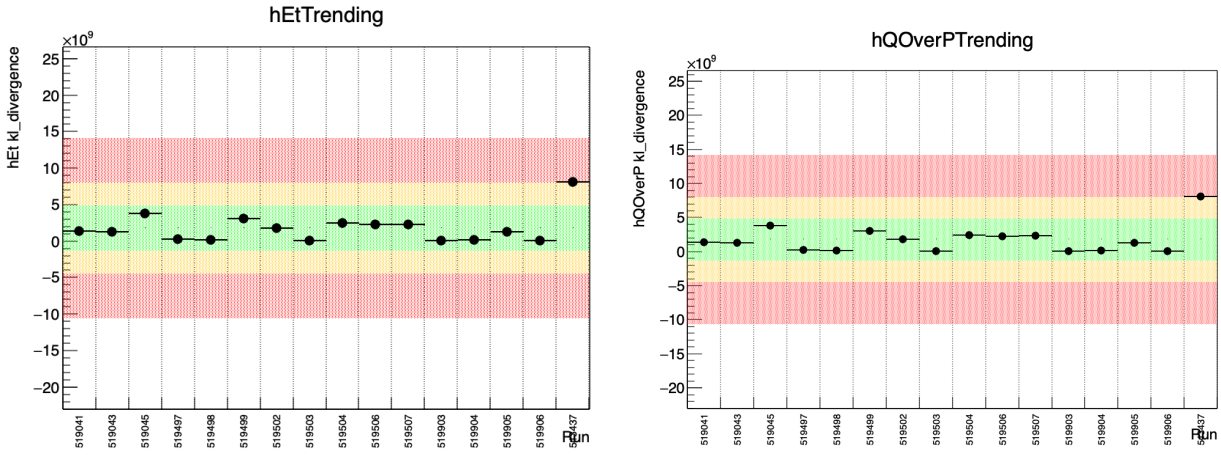

Figure 2.9: KL divergence calculations and quality ratings for η and Charge/Transverse momentum for each run.

5. **Outlier Detection:** Identify outliers or significant shape distortions by analyzing the KL divergence values. Histograms with high KL divergence values are flagged as potential outliers, indicating deviations from the expected data quality.

2.5.3 Results and Discussion

By flagging these outliers, the method enabled prompt identification and correction of data quality issues, ensuring that only high-quality data were used for subsequent analysis. This approach proved to be an effective tool for maintaining the integrity of the data collected by the TPC.

The use of KL divergence for data quality assessment in the TPC QC group has provided a robust and systematic method for identifying outliers and significant shape distortions in histograms from different runs of data. By comparing each individual histogram to an average histogram, this method ensures that deviations from expected distributions are promptly detected and addressed. This work highlights the importance of statistical measures in maintaining data quality in high-energy physics experiments.

This method, using the language of information theory, provides a practical example of how entropy can be used to quantify the degree of correlation between two distributions.

Chapter 3

Measuring the Final State using ALICE Data

As described in the first chapter, information in the the final state is contained in the distribution of particles produced in p-p collisions. This distribution, known as the multiplicity distribution, describes the probability of finding N primary produced particles. This distribution will be extracted from data taken by the ALICE collaboration at CERN. In this chapter we will explain how data is taken, corrected, and then analysed to create the final-state multiplicity distribution. Distributions have already been published for lower energies at the LHC up to 8 TeV [62], this chapter describes the procedure used to measure multiplicity at the next highest energy of 13 TeV.

3.1 Simulated and Real Data Sets Used for 13 TeV

The following ALICE data sets were used for generating the multiplicity distribution at 13 TeV and error estimation.

Data Type	**Period**	**Number of Events**
Data	LHC15f	28.9 M
MC (PYTHIA)	LHC15g3c3	3.2 M
MC (EPOS)	LHC16d3	12.8 M

Table 3.1: Summary of data periods and number of events after cuts for INEL collisions

Runs
226062, 226170, 226220, 226225, 226444, 226445, 226452,
226466, 226468, 226472, 226476, 226483, 226495, 226500

Table 3.2: List of runs

3.2 Kinematic Variables

Because we are making a comparison of experimental data in the final-state to quantities defined by the initial-state, it is necessary to relate the kinematics of the final-state to the kinematics that define the initial-state; namely the Bjorken-x and the energy scale Q^2.

3.2.1 Rapidity and Pseudorapidity

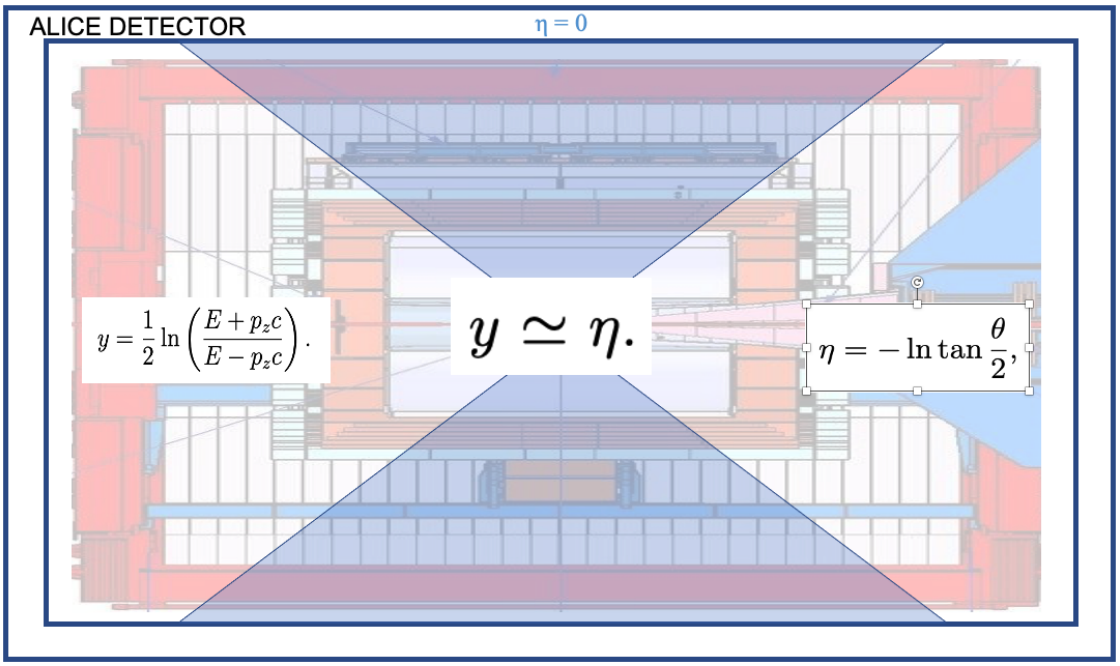

Figure 3.1: Psuedorapidity (η) is a geometric angle in relation to the beam axis. For low mass particles it can be shown that η is equal to rapidity.

Rapidity (y) and pseudorapidity (η) are both measures used in high-energy particle physics to describe the angle of a particle relative to the beam axis. Rapidity is a Lorentz invariant quantitiy defined in terms of the particle's energy (E) and the longitudinal momentum component (p_z) as:

$$y = \frac{1}{2}\ln\left(\frac{E+p_zc}{E-p_zc}\right). \tag{3.1}$$

Pseudorapidity, on the other hand, is defined based on the polar angle (θ) of the particle with respect to the beam axis:

$$\eta = -\ln\tan\left(\frac{\theta}{2}\right). \tag{3.2}$$

In the limit where the particle's mass is negligible compared to its momentum (i.e., for ultra-relativistic particles), rapidity approximates pseudorapidity ($y \approx \eta$). This approximation

simplifies analysis and calculations in collider experiments, as pseudorapidity depends only on the angle θ and is easier to measure experimentally. The relationship between rapidity and pseudorapidity allows physicists to use pseudorapidity as a practical tool for describing particle trajectories in detectors, particularly when dealing with high-energy collisions where particles are produced with large momenta. In p-p collisions the majority of produced hadrons is constituted by low-mass particles that allow for a neat perfect correspondence between rapidity and pseudorapidity.

3.2.2 Mapping Final-State Kinematics to Initial-State Kinematics

The initial state of partons is defined in terms of the Bjorken-x which is defined as the fraction of the protons momentum carried by a struck parton. The magnitude of this momentum is related to the angle of the particle produced in the interaction. In order to make a mapping of the initial-state kinematics to the final-state kinematics we relate x to the pseudorapidity; the derivation goes as follows:

$$y_p = \frac{1}{2} \ln \left(\frac{E_p + p_{zp}}{E_p - p_{zp}} \right) \tag{3.3}$$

$$y_h = \frac{1}{2} \ln \left(\frac{E_h + p_{zh}}{E_h - p_{zh}} \right) \tag{3.4}$$

$$q = P - P_h \tag{3.5}$$

$$x = -\frac{q^2}{2P \cdot q} \tag{3.6}$$

$$P \cdot q = P \cdot (P - P_h) = (P \cdot P) - (P \cdot P_h) \tag{3.7}$$

$$x = -\frac{(P_h \cdot P)}{(P \cdot P)} \tag{3.8}$$

$$\frac{p_{zh}}{p_{zp}} = \frac{(P_h \cdot P)}{(P \cdot P)} \tag{3.9}$$

$$x = \frac{1}{\frac{p_{zh}}{p_{zp}}} \tag{3.10}$$

$$\ln(1/x) = \ln\left(\frac{p_{zh}}{p_{zp}}\right) = \ln\left(\frac{E_h + p_{zh}}{E_h - p_{zh}}\right) - \ln\left(\frac{E_p + p_{zp}}{E_p - p_{zp}}\right) \tag{3.11}$$

$$\ln(1/x) = y_p - y_h \tag{3.12}$$

Using the relation between rapidity and x one can create a distribution of the average value of x in a p-p collision. The mean and limits of this distribution will define the scale and range over which we compare to the initial state.

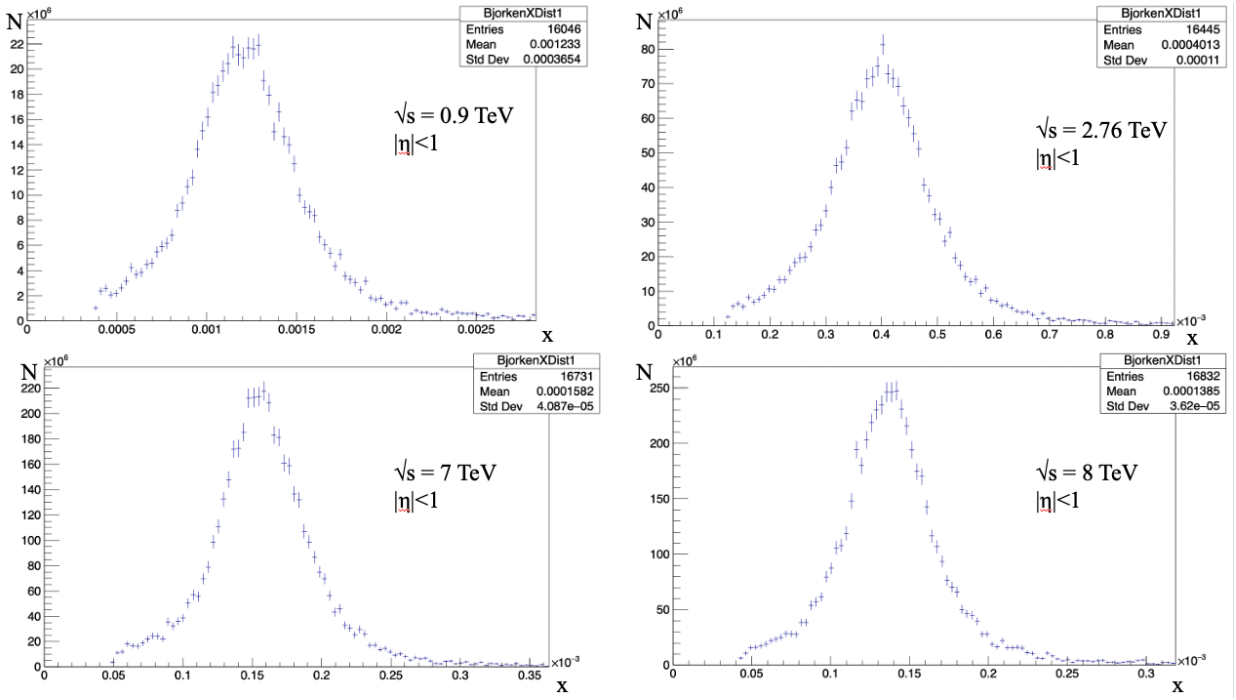

Figure 3.2: Distribution of x calculated from η from PYTHIA for different energies.

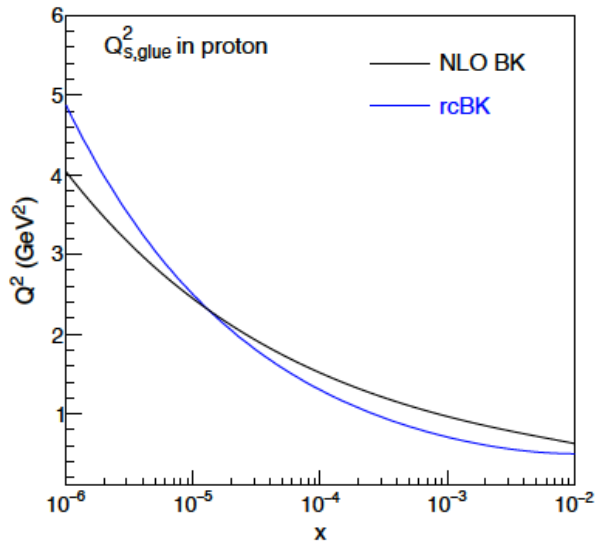

Figure 3.3: Gluon saturation scale defined by NLO BK calculations [64]

Another variable which defines the initial state is the energy scale at which the proton is probed. In high energy hadronic collisions the appropriate scaling variable is known as

the saturation scale. The relationship between the saturation scale and the Bjorken-x is characterized by the scaling behavior $Q_s^2(x) \sim x^{-\lambda}$, where λ is a parameter that depends on the QCD dynamics and can be extracted from experimental data using the Balitsky-Kovchegov (BK) equations [64].

3.3 Data Selection

To ensure high quality data is used in track counting quality cuts are made on both the events and the individual reconstructed tracks. These cuts represent standard cuts done for multiplicity studies performed in ALICE [62].

3.3.1 Event Selection

Event selection criteria must be set to remove events with low quality reconstruction or events that are significantly impacted by the detector. A crucial part of this process was applying a cut to the primary vertex in the z-component. This cut removed events in which the z-component (along the beam axis) of the primary vertex was outside -10 cm and 10 cm. This criterion is important because it excludes events that occurred far from detectors geometric center, reducing the impact of detector inefficiencies and biases introduced by non-uniform acceptance and efficiency in the forward and backward rapidity regions.

Multiplicity distributions are typically given in 3 different event classes requiring different trigger conditions:

1. INEL: Includes all inelastic events. The trigger conditions required the presence of a hit in either the Silicon Pixel Detector (SPD), the V0A, or the V0C.

2. INEL>0: Includes all inelastic events where at lest 1 charged particle was produced in $-1 < \eta < 1$. Trigger condition is the same as for the INEL event class.

3. NSD: Includes inelastic events other than single-diffractive events. The trigger conditions required the presence of a hit in the Silicon Pixel Detector (SPD), the V0A, and the V0C.

3.3.2 Pile-up rejection

Different background sources can influence the measurement of the multiplicity distribution, a key issue being pile-up events. Pile-up is a phenomenon that occurs when multiple proton-proton interactions from the same bunch-crossing are detected as one event. This effect is significant in high-luminosity environments such as those found at the LHC, where a single bunch crossing can lead to multiple independent collisions. The AliAnalysisUtils class was used to identify and reject pileup. AliAnalysisUtils is a utility class that provides various functionalities. The functionalities used here pertain to pile-up event selection and rejection. The following cuts were applied:

1. **util.SetMinPlpContribMV(5):** This sets the minimum number of contributors to the pile-up vertex in the multi-vertex pile-up rejection method. Essentially, it is a parameter that defines the minimum number of particles that must contribute to a collision for it to be considered a pile-up event.

2. **util.SetMaxPlpChi2MV(5):** Sets the maximum χ^2 value for the pile-up vertex in the multi-vertex pile-up rejection method. Chi-squared is a statistical measure used to determine the goodness of fit of an observed distribution to a theoretical one.

3. **util.SetMinWDistMV(15):** This sets the minimum "weighted distance" between primary and pile-up vertices in the multi-vertex pile-up rejection method. Essentially, it is a parameter that defines the minimum spatial distance that must exist between two collision events for them to be considered distinct (and thus for one to be considered

a pile-up event).

3.3.3 Quality Cuts on Tracks

In high-energy physics experiments, it is crucial to apply standard cuts on tracks to ensure the quality and reliability of the data used for creating multiplicity distributions. These cuts help in rejecting noise and poorly reconstructed tracks, thereby improving the precision of the measurements. The following criteria are some of the cuts commonly used for track selection in ALICE:

- **ITS hits:** For a track to be included in the multiplicity analysis, it must have hits in at least 2 layers of the ITS. This ensures that the track is well-defined and has a reliable position measurement near the primary vertex.

- **TPC clusters:** To ensure the accuracy of the track reconstruction, a minimum number of TPC clusters (hits) is required. A track must have at least 70 TPC clusters out of a possible 159 to be considered a good-quality track. This requirement helps in reducing the contribution of short, poorly reconstructed tracks.

- **Track Fit Quality:** Tracks must have a reduced chi-square (χ^2/NDF) value for the TPC track fit less than 4. This criterion ensures that the track fit is reasonable and the reconstructed trajectory is consistent with the measured points.

- **Track Length:** Tracks must have a minimum length in the TPC to ensure they are not spurious or short tracks. Typically, a minimum track length of 120 cm is required.

3.4 Data corrections

Detector effects can significantly smear the final measured multiplicity distribution in high-energy physics experiments. These effects arise from various sources such as finite detector resolution, inefficiencies in track reconstruction, and noise. Finite resolution can cause inaccuracies in the determination of particle positions and momenta, leading to errors in counting the actual number of particles. Inefficiencies in track reconstruction, where some tracks are missed or falsely identified, can lead to an underestimation or overestimation of the true multiplicity. Additionally, background noise and secondary interactions can introduce spurious tracks, further distorting the multiplicity distribution. To determine these effects Monte-Carlo simulations of the detector mechanics and geometry are used in conjunction with simulated physics to resolve the detectors influence on the measurement. Detectors are typically simulated using a software package known as GEANT [65].

3.4.1 Bayesian Analysis

A commonly used method for determining and correcting for detector effects is Bayesian Unfolding. This unfolding procedure adopts a Bayesian statistical approach, iteratively refining the estimate of the true distribution.

Given a prior distribution, an unfolding matrix is computed and used to obtain an unfolded distribution. The latter then serves as a new prior for the next iteration. This process repeats until the method converges to a stable solution.

Below are the steps followed in the Bayesian unfolding procedure [62, 66]:

1. **Definition of Response Matrix:** The response matrix, R_{ij}, defines the probability that a true event in bin i is observed in bin j. This matrix is determined from Monte Carlo simulations of the detector response (determined by GEANT) to known input

distributions (determined by event generators like PYTHIA).

$$R_{ij} = \frac{\text{Number of events generated in bin } i \text{ and reconstructed in bin } j}{\text{Total number of events generated in bin } i} \tag{3.13}$$

2. **Initialization:** The procedure begins with an initial prior, $P^{(0)}$, which is chosen to be the measured distribution, M.

$$P^{(0)} = M \tag{3.14}$$

3. **Computation of Unfolding Matrix:** For the n-th iteration, an unfolding matrix $M^{(n)}$ is computed as follows:

$$M_{ji}^{(n)} = \frac{R_{ij} P_i^{(n-1)}}{\sum_k R_{kj} P_k^{(n-1)}} \tag{3.15}$$

where R_{ij} denotes the element of the response matrix in the i-th row and j-th column.

4. **Unfolding:** The unfolded distribution, $U^{(n)}$, is then calculated by applying the unfolding matrix $M^{(n)}$ to the measured distribution, M.

$$U_i^{(n)} = \sum_j M_{ji}^{(n)} M_j \tag{3.16}$$

5. **Update of Prior:** The unfolded distribution $U^{(n)}$ is normalized and serves as the new prior $P^{(n)}$ for the next iteration.

$$P^{(n)} = \frac{U^{(n)}}{\sum_i U_i^{(n)}} \tag{3.17}$$

This iterative process continues until a predetermined number of iterations is reached. We find the optimal number of iterations by observing the (χ^2) value.

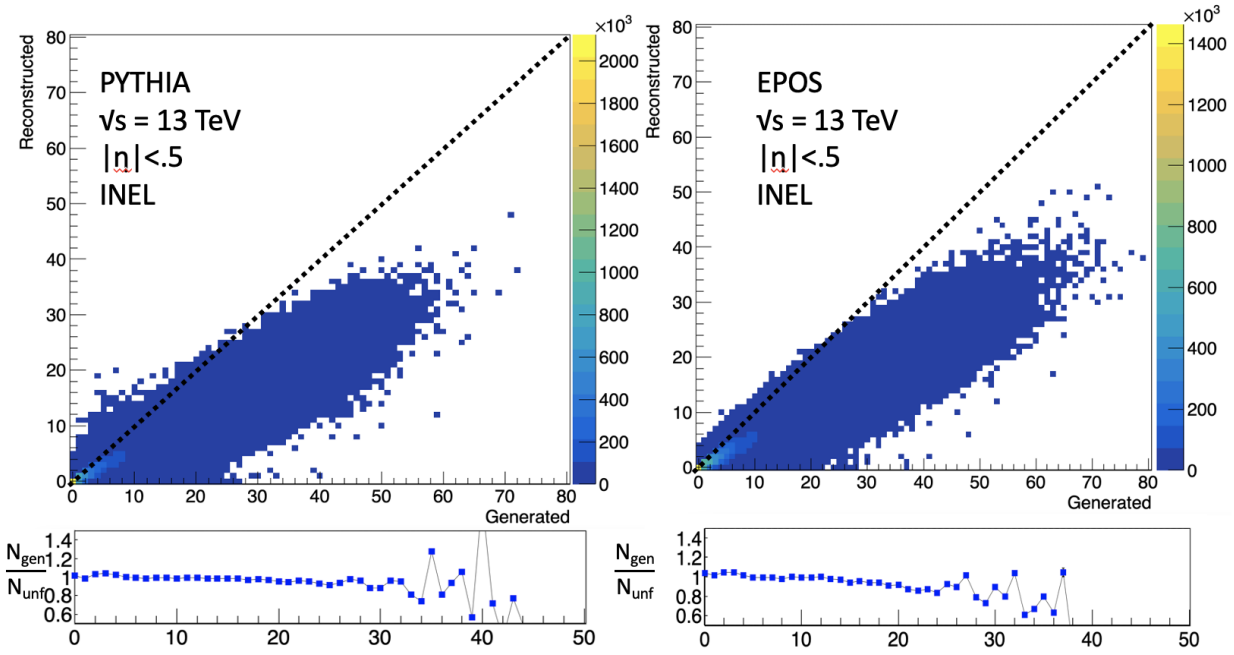

Figure 3.4: Response matrices for INEL events using PYTHIA and EPOS event generators. Ratios compare PYTHIA generated true distributions to unfolded distributions.

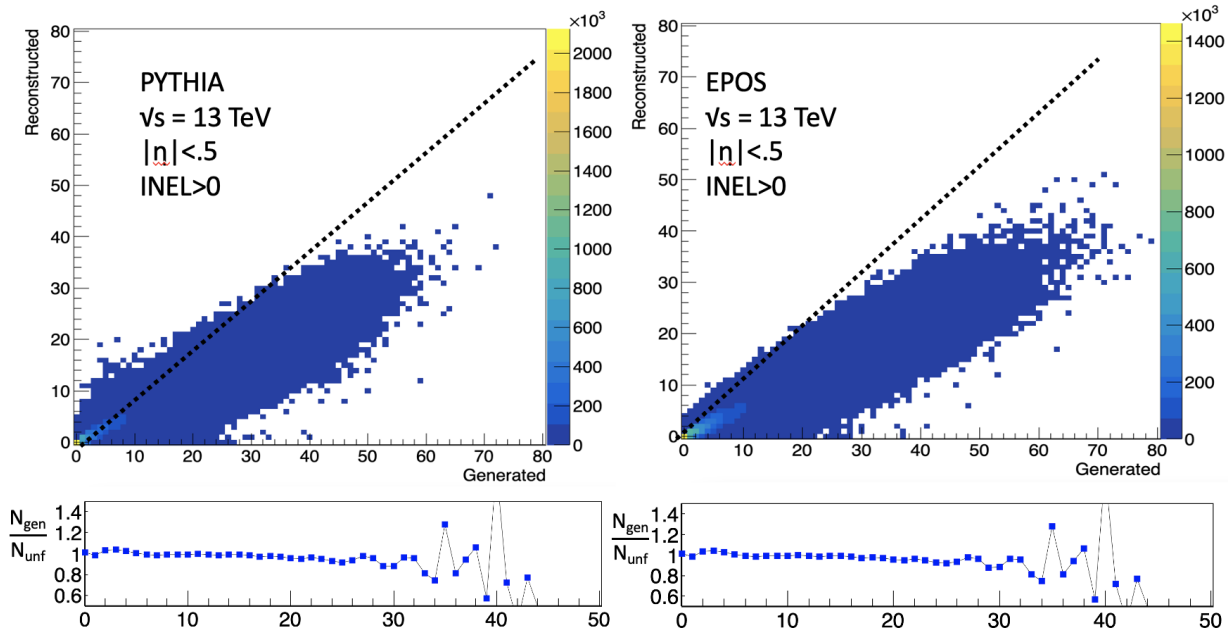

Figure 3.5: Response matrices for INEL>0 events (INEL events with at least one particle generated at mid rapidity) using PYTHIA and EPOS event generators. Ratios compare PYTHIA generated true distributions to unfolded distributions.

χ^2 Convergence in Bayesian Unfolding

The iterative Bayesian unfolding process relies on a metric to determine the convergence of the solution. A common choice for this metric is the chi-square (χ^2) statistic, which quantifies the deviation of the unfolded distribution from the true distribution, normalized by the statistical uncertainty.

The χ^2 value for the n-th iteration is computed as follows:

$$\chi^2_n = \frac{1}{N} \sum_i \frac{(U_i^{(n)} - T_i)^2}{\sigma^2_{T_i}} \tag{3.18}$$

where N is the number of bins, T_i represents the true distribution, and σ_{T_i} is the statistical uncertainty associated with T_i.

For each iteration, the difference between the unfolded and true distributions is calculated

for each bin and then squared. This squared difference is divided by the variance of the true distribution to normalize it, taking into account the associated uncertainties. Summing up these normalized squared differences for all bins and then dividing by the number of bins gives the χ^2 value.

A lower χ^2 value signifies a closer match between the unfolded and true distributions. The procedure aims to minimize this statistic to obtain the best possible estimate of the true distribution.

The unfolding process is considered to have converged when the χ^2 value becomes stable, i.e., when the change in χ^2 from one iteration to the next falls below a predefined threshold. At this point, the unfolded distribution is regarded as the best estimate of the true distribution, given the measured data and response matrix.

The significance of the χ^2 statistic lies in its role as a goodness-of-fit indicator, providing a quantifiable measure of how well the unfolded distribution approximates the true distribution, given the inherent statistical uncertainties.

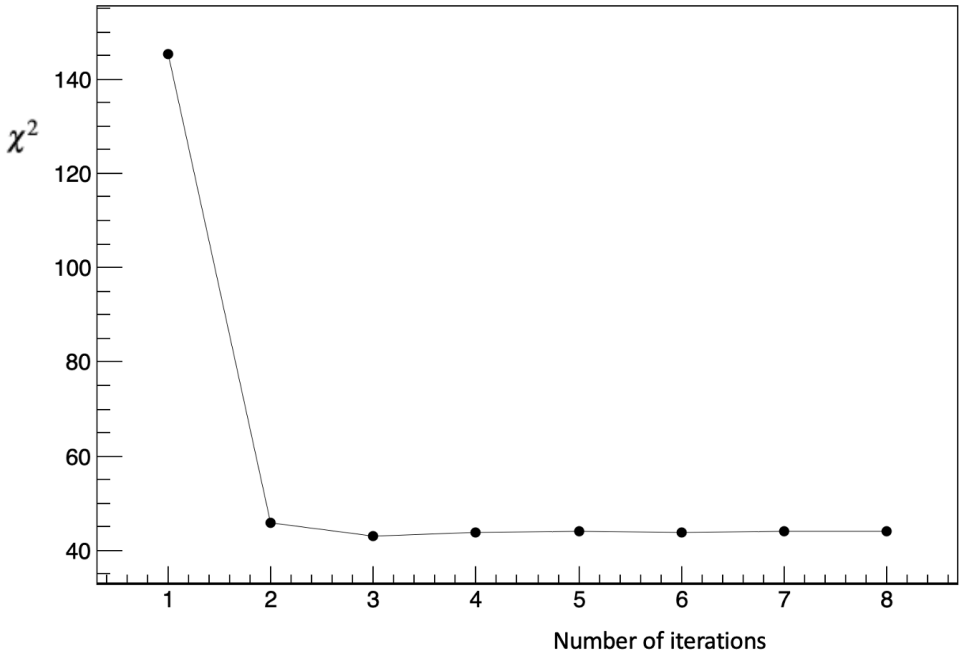

Figure 3.6: χ^2 vs number of iterations, the unfolding procedure seeks to minimize this value until a stable point is reached.

Closure test

To ensure our unfolding procedure is correctly producing the true distribution a closure test is performed. This is done by splitting a MC produced set into 2 subsets. One subset is used to generate the response matrix, the other is used to produce a generated and reconstructed distribution. The two MC data sets both use all the same tunes to simulate the detector response and the same event generator (PYTHIA) to simulate the collision. If the unfolding procedure is working we should be able to unfold the reconstructed distribution from subset one using the response matrix obtained from subset two, and recover the generated distribution of subset one.

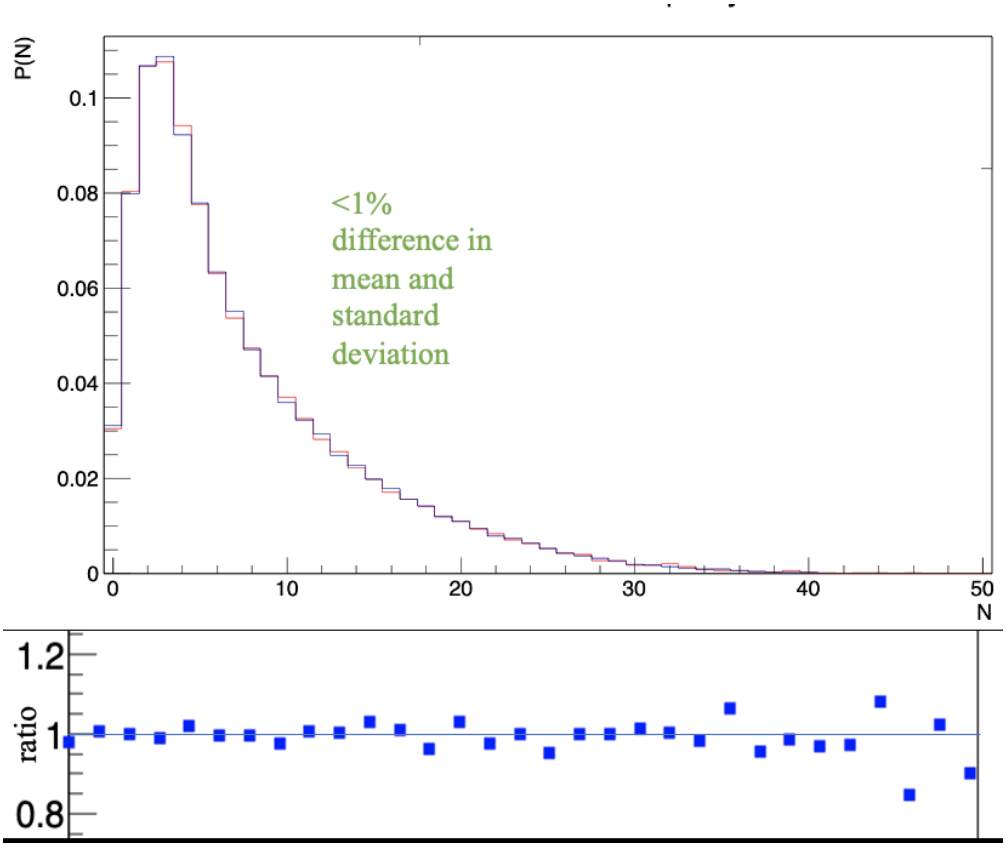

Figure 3.7: Generated and unfolded MC distributions to demonstrate the effectivness of the unfolding procedure. The ratio should be near 1, deviations from one contribute to the MC non-closure systematic error

Corrected Results

After performing the unfolding procedure we have the final distributions corrected for detector effects. We observe an increase in average multiplicity after the correction. This increase in the average multiplicity is expected because the detector response is dominated by inefficiencies and tracks are lost in the measuring process.

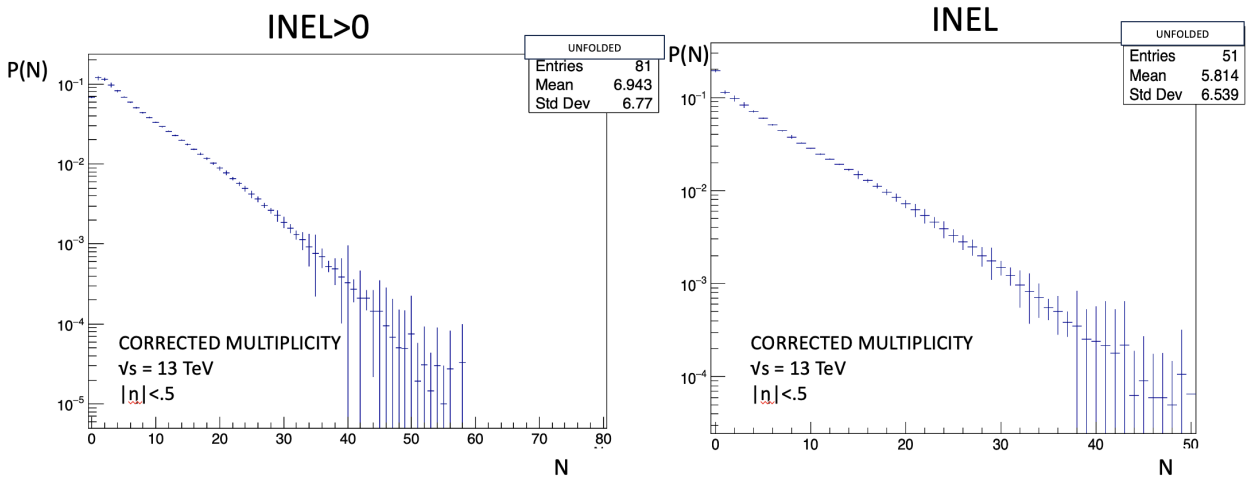

Figure 3.8: Corrected multiplicity distribution for 13 TeV.

3.5 Error Estimation

Error estimation is crucial in data analysis to quantify the reliability and precision of the measured results. Statistical errors arise from the inherent randomness and limited size of the data sample, and they can be reduced by increasing the sample size. Systematic errors, on the other hand, stem from biases or inaccuracies in the measurement process or experimental setup, and they are not reduced by increasing the sample size. Both types of errors should be carefully estimated and considered to ensure that the final results accurately reflect the true values and to account for any potential biases or uncertainties in the measurements. Proper error estimation allows for meaningful comparisons between experimental data and theoretical predictions, and it is essential for drawing reliable scientific conclusions.

3.5.1 Statistical Error

The statistical uncertainties of the unfolded distribution are directly obtained from the diagonal elements of the covariance matrix. For a given variable, the variance is defined as the expectation of the squared deviation of this variable from its expected value. Thus, the

square root of the diagonal elements of the covariance matrix provides the standard deviation, or the statistical uncertainty, associated with each bin in the unfolded distribution.

The covariance matrix Σ is given by:

$$\Sigma = (R^T V^{-1} R)^{-1}, \tag{3.19}$$

where R is the response matrix, V is the covariance matrix of the measured distribution, and R^T is the transpose of the response matrix. Here, V^{-1} is the inverse of the covariance matrix of the measured distribution, which weights the response matrix in accordance with the measurement uncertainties.

The variance (square of the statistical uncertainty) for the i^{th} bin in the unfolded distribution is then the i^{th} diagonal element of the covariance matrix Σ, denoted as Σ_{ii}:

$$\mathrm{Var}(D^{(n)}_{\mathrm{unfolded},i}) = \Sigma_{ii}, \tag{3.20}$$

And the statistical uncertainty for the i^{th} bin is the square root of the variance:

$$\sigma_{D^{(n)}_{\mathrm{unfolded},i}} = \sqrt{\Sigma_{ii}}. \tag{3.21}$$

Thus, the covariance matrix not only provides the uncertainties of the unfolded distribution but also information about how these uncertainties are correlated across different bins.

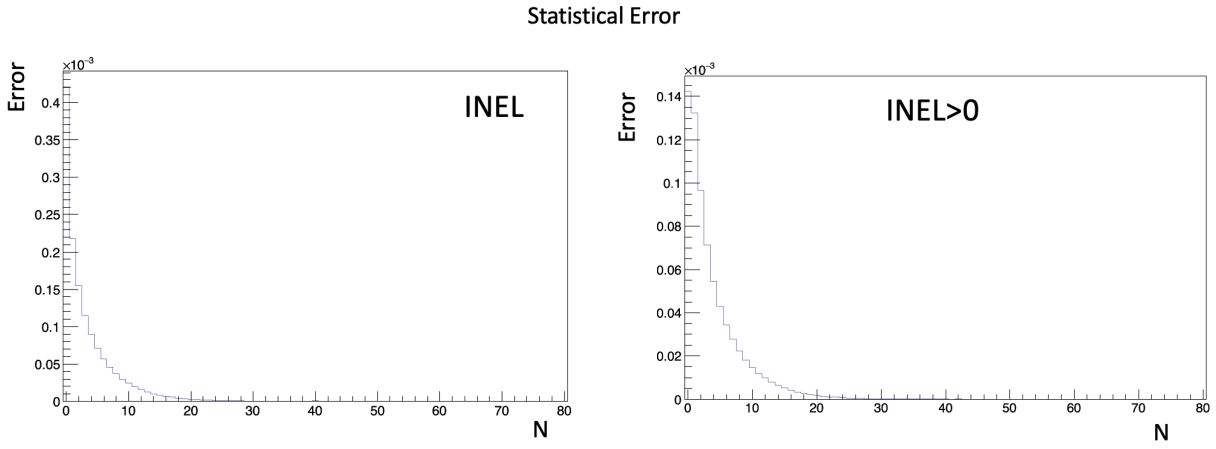

Figure 3.9: Statistical errors calculated from the diagonal of the unfolded covariant matrix

3.5.2 Systematic Error

In this section I will describe a few sources of systematic error and how they contribute to the overall error estimation.

Monte-Carlo non-closure systematic uncertainty

The Monte Carlo non-closure uncertainty is a type of systematic uncertainty that arises when we estimate physical quantities by unfolding detector-level measurements to the particle or parton level. Unfolding is a technique used to correct the smeared and distorted data due to the finite resolution and efficiency of the detector, and to account for the acceptance. To estimate this uncertainty, we can use a simulated Monte Carlo dataset, generated from the PYTHIA event generator.

The procedure is as follows [62]. Firstly, generate a 'truth' distribution D_{gen} and a 'reconstructed' distribution D_{rec} from the PYTHIA Monte Carlo data, where D_{rec} has been processed through the detector simulation. Then using a separate subset of the same MC data generate the response matrix. Next, perform the unfolding procedure on the 'reconstructed' distribution D_{rec} to obtain an 'unfolded' distribution D_{unfolded}. The ratio $r = \frac{D_{\text{gen}}}{D_{\text{unfolded}}}$ of the

generated or true distribution to the unfolded distribution is then calculated. In an ideal situation, this ratio should be unity everywhere. However, due to statistical fluctuations and potential biases in the unfolding procedure, there may be deviations from unity. These deviations are then used to estimate the non-closure uncertainty. The relative non-closure uncertainty for each bin can be calculated as $|1 - r|$. For the absolute uncertainty, one can then multiply this relative uncertainty by the value of the bin content itself.

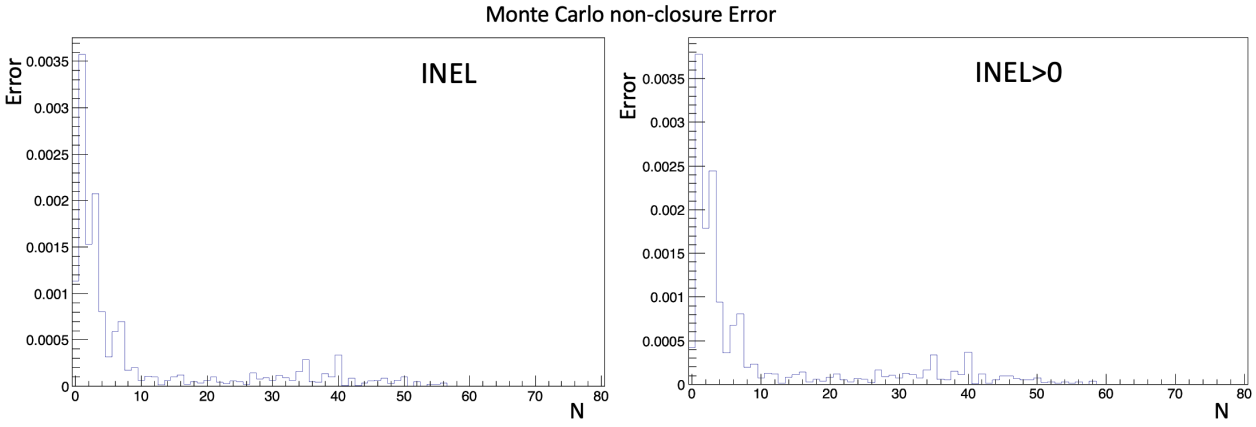

Figure 3.10: Relative systematic errors due to the performance of the unfolding procedure.

Event generator systematic uncertainty

Systematic uncertainties can arise from the choice of the event generator used in the simulation and unfolding process. This is due to the fact that different event generators may model the underlying physics of the collision events differently, resulting in different response matrices and thus different unfolded results. To quantify this uncertainty, one can compare the unfolded results obtained from different event generators [62].

In this case, we have two histograms: one is reconstructed using PYTHIA and unfolded using a response matrix derived from PYTHIA, let's denote this as $D_{\text{unfolded, PYTHIA}}$; the other is reconstructed using PYTHIA and unfolded using a response matrix derived from EPOS [67](another event generator), denote this as $D_{\text{unfolded, EPOS}}$. The ratio $r = \frac{D_{\text{unfolded, EPOS}}}{D_{\text{unfolded, PYTHIA}}}$

can then be calculated.

If the two event generators model the underlying physics identically, this ratio should be unity. However, due to the different approximations and physics inputs in the two generators, there may be deviations from unity. These deviations give an estimate of the systematic uncertainty due to the choice of the event generator. The relative event generator systematic uncertainty for each bin can be calculated as $|1-r|$. To obtain the absolute uncertainty, one can multiply this relative uncertainty by the value of the bin content itself.

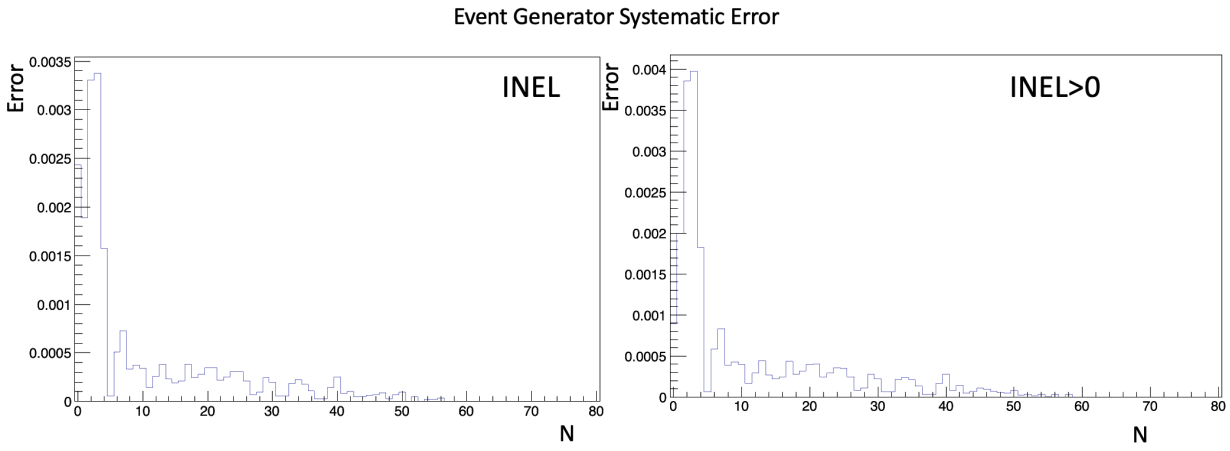

Figure 3.11: Relative systematic errors due to event generator dependence of the unfolding procedure.

3.5.3 Total Error

Finally once each source of error is determined and quantified the total error is obtained by summing each of these sources of error per bin. Comparing each source of error it is clear that the statistical error is negligible due to the high number of statistics provided by the LHC. Systematic errors associated with non-closure in the unfolding procedure and choice of event generator are small but not negligible. These systematic errors are particularly high for low multiplicity events, as shown in the figure below.

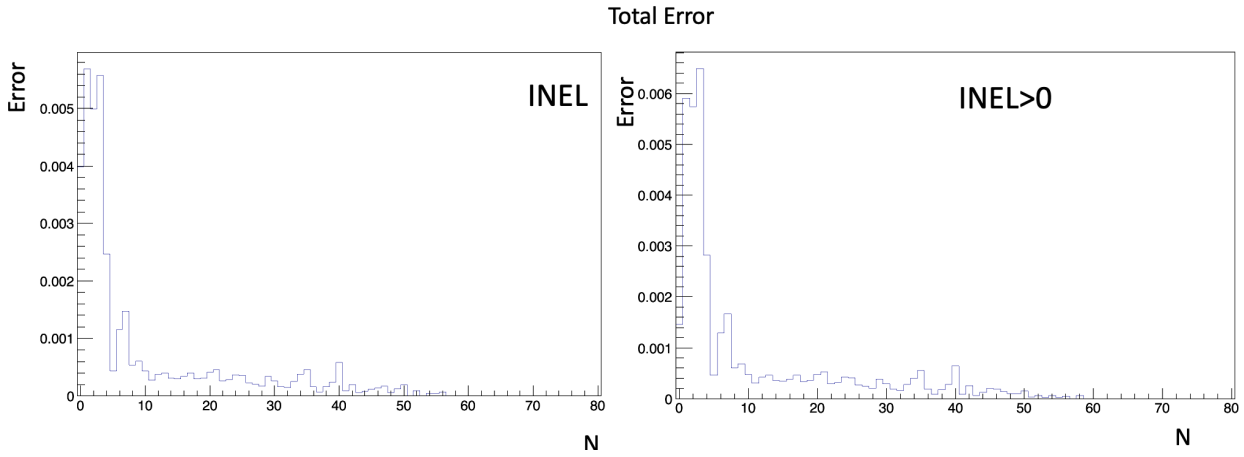

Figure 3.12: Total relative (percent) errors for each bin the multiplicity distribution including statistical errors and both sources of systematic error.

3.6 Final State Entropy

Entropy in the final state is defined by the level of disorder in the multiplicity distribution P(n). P(n) defines the probability of finding n produced hadrons after the collision. We can quantify the level of disorder in this distribution using the Shannon entropy. This will by calculated using multiplicity distributions from published results up to 8 TeV and the results of the measurement at 13 TeV described in the previous chapter. A comparison of the final-state entropy will also be made between data and PYTHIA using different tunes.

3.6.1 Mean Halving Transformation

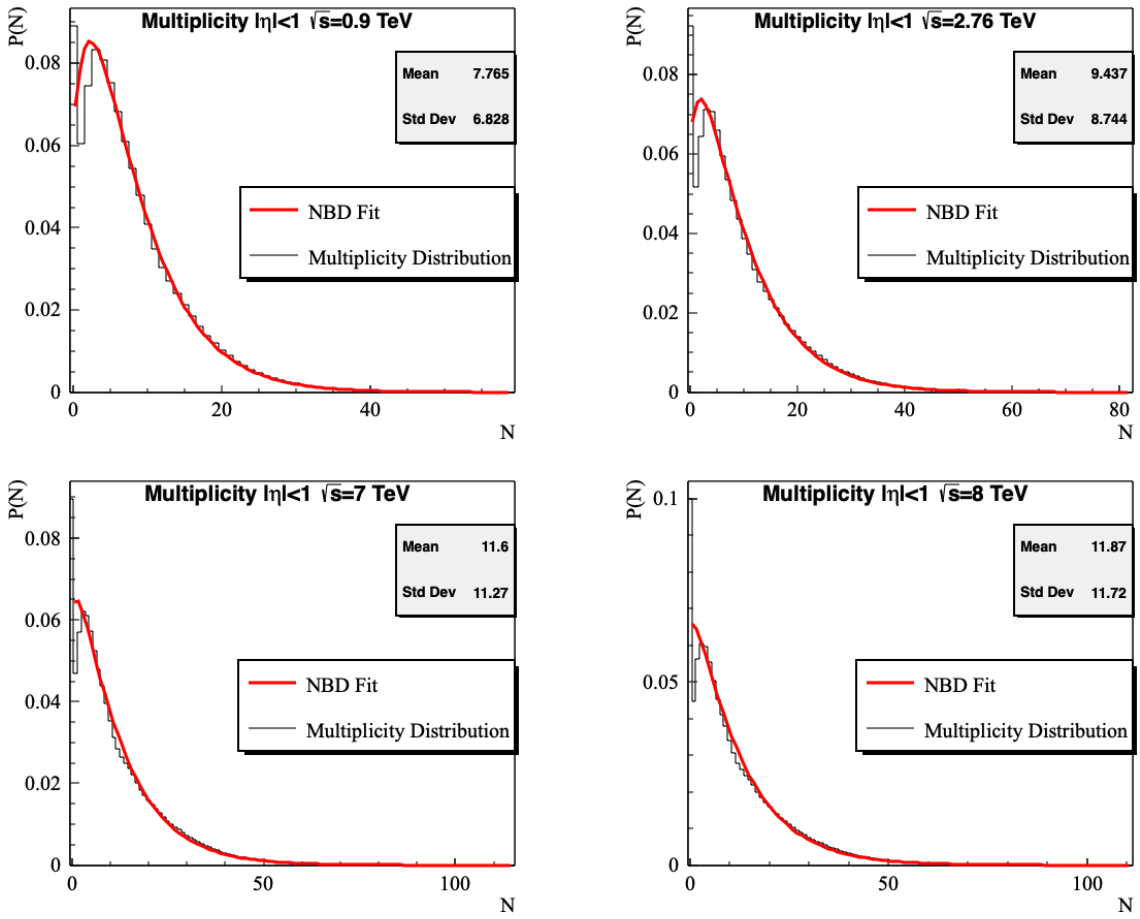

Figure 3.13: Produced hadron multiplicity distributions from ALICE data [62]

Above are multiplicity distributions from p-p collisions measured at $|\eta| < 1$. These distributions P(N) represent the probability of finding N particles within that region. Because these distributions are generated by both protons a transformation on the distribution needs to be made in order to make a meaningful comparison to PDFs, which represent the parton distribution in a single proton. In making this transformation an assumption is made that half of the produced hadrons are coming from one proton and half is coming from the other. Under this assumption we "squeeze" the distribution such that the mean is reduced by half. We will describe two equivalent methods for performing this transformation, one involving

a fit and one using a bin counting method.

Negative Binomial Distribution Fit

The negative binomial distribution (NBD) is a discrete probability distribution that is widely used to describe the multiplicity distributions of particles produced in proton-proton (p-p) collisions.

Mathematically, the probability mass function of the negative binomial distribution is given by [62]:

$$P_{\mathrm{NBD}}(n;\langle n\rangle,k)=\frac{\Gamma(n+k)}{\Gamma(k)\Gamma(n+1)}\left(\frac{\langle n\rangle}{\langle n\rangle+k}\right)^{n}\left(\frac{k}{\langle n\rangle+k}\right)^{k}, \tag{3.22}$$

where n is the number of particles produced, $\langle n\rangle$ is the average number of particles (mean multiplicity), k is a parameter related to the dispersion of the distribution, and Γ denotes the gamma function.

The parameter k controls the shape of the distribution. When k is large, the negative binomial distribution approaches the Poisson distribution, which appears to be the case as we increase the collision energy.

In the context of p-p collisions, the NBD provides a good fit to the observed multiplicity distributions across a wide range of energies. It captures features of the data, such as the long tail at high multiplicities, which is indicative of events with higher particle production. It should be noted that in order to capture the shape of the distribution a low multiplicity a double weighted NBD is required.

The method for transforming the distribution using this fit was described by Tu et. al [51]. The procedure involves taking the NBD fit to the distribution and cutting the parameter associated with the mean in half while keeping the remainder of the parameters the same. Essentially the mean is cut in half while the relative variance is kept constant.

Bin Counting

A simpler method for performing the mean halving transformation involves bin counting. In this method, a new histogram is created where each bin is filled with the value from the corresponding bin in the original histogram, but associated with twice the number of particles. Mathematically, if $P_{\text{original}}(N)$ represents the original multiplicity distribution, then the new distribution $P_{\text{new}}(N)$ is defined as:

$$P_{\text{new}}(N) = P_{\text{original}}(2N). \tag{3.23}$$

This means that the value of the new distribution at bin N is taken from the original distribution at bin $2N$. For example, the bin associated with $N = 1$ in the new distribution will be filled with the value from the original distribution at $N = 2$ such that $P_{\text{new}}(1) = P_{\text{original}}(2)$, and so on.

By applying this transformation, the mean of the multiplicity distribution is effectively halved, which allows for a meaningful comparison to parton distribution functions (PDFs) that represent the parton distribution in a single proton.

Final Distribution

Once the transformation has been made the distributions are re-normalized to unity so that they are reflective of a true probability distribution. A comparison of the distribution after transforming using both methods is shown below. It is clear from the figure that both methods are relatively equivalent, and will give very similar values for the final state entropy. In this analysis the distribution transformed with our bin counting method was used for entropy calculations.

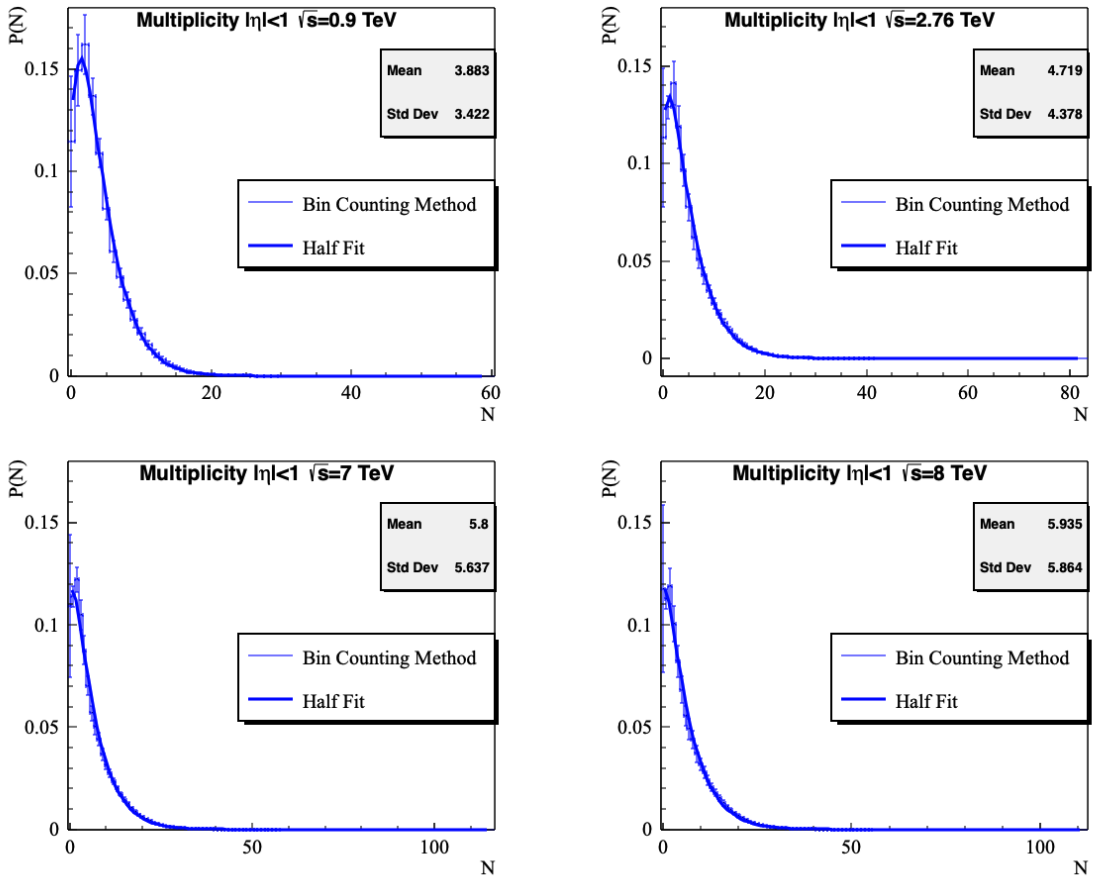

Figure 3.14: Transformed multiplicity distribution reflective of a single proton using a fit method and a bin counting method.

3.6.2 Comparison to PYTHIA

Now with corrected data and a distribution reflective of a single proton we can quantify the level of disorder in the system by summing over the distribution in the following way:

$$S(N) = -\sum_{n=1}^{\infty} p_n \ln p_n, \tag{3.24}$$

Where S(N) represents the Shannon entropy in the particle number basis[1], and p_n represents the probability of the nth bin.

[1] defined in chapter 1

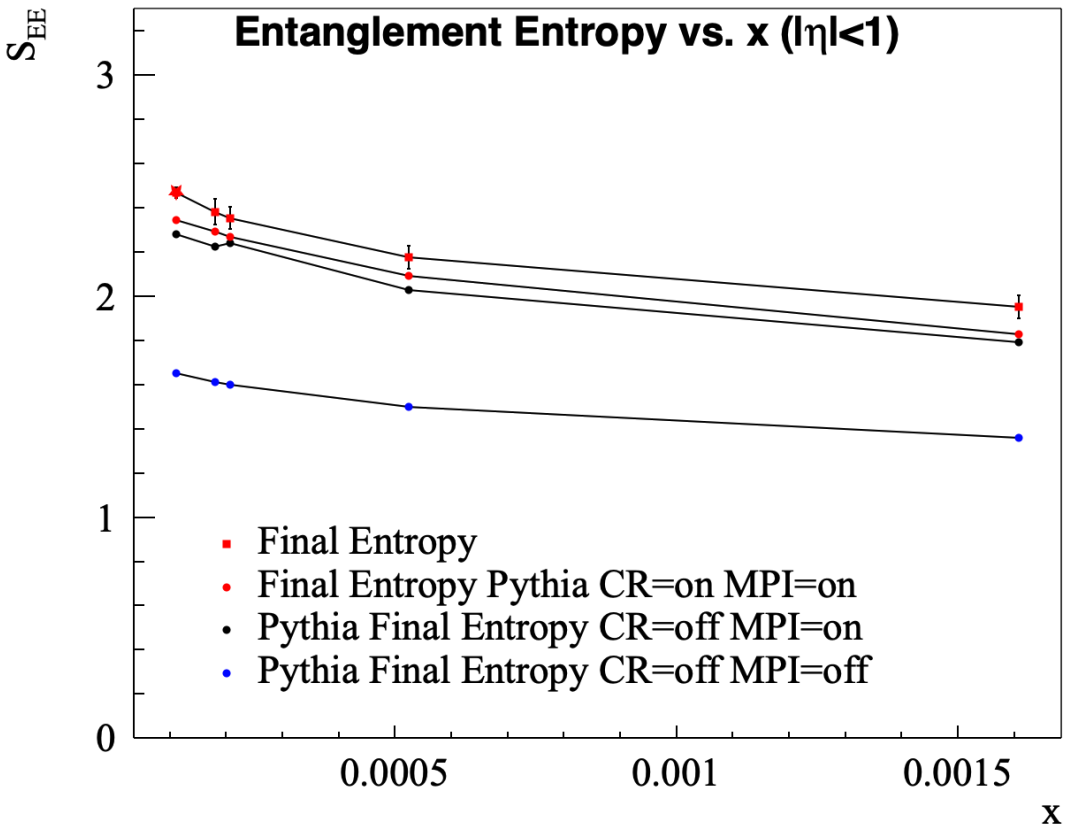

Figure 3.15: Final state entropy compared to the event generator PYTHIA using 3 diffrent tunes.

Entropy is calculated using data from ALICE and compared to entropy calculated from distributions generated by PYTHIA. Three different tunes of PYTHIA were used[2] either allowing or disallowing CR and MPI. The color reconnection model was tuned considering both two string recombination modes (leading color) and three string color recombination modes (beyond the leading color).

It is clear that PYTHIA underestimates the final state entropy even with the most sophisticated tune involving both MPI and CR. PYTHIA does not take into account entanglement of the initial state partons, this may be a possible explanation for why the model underestimates the data. Furthermore, when MPI is off the entropy is vastly underestimated and the

[2]MPI and CR are defined in the chapter 1 in the section on PYTHIA

slope of the points is reduced at low-x. It may be possible that the MPI tune is mimicking at least partially the effects of entanglement.

Chapter 4

Initial State

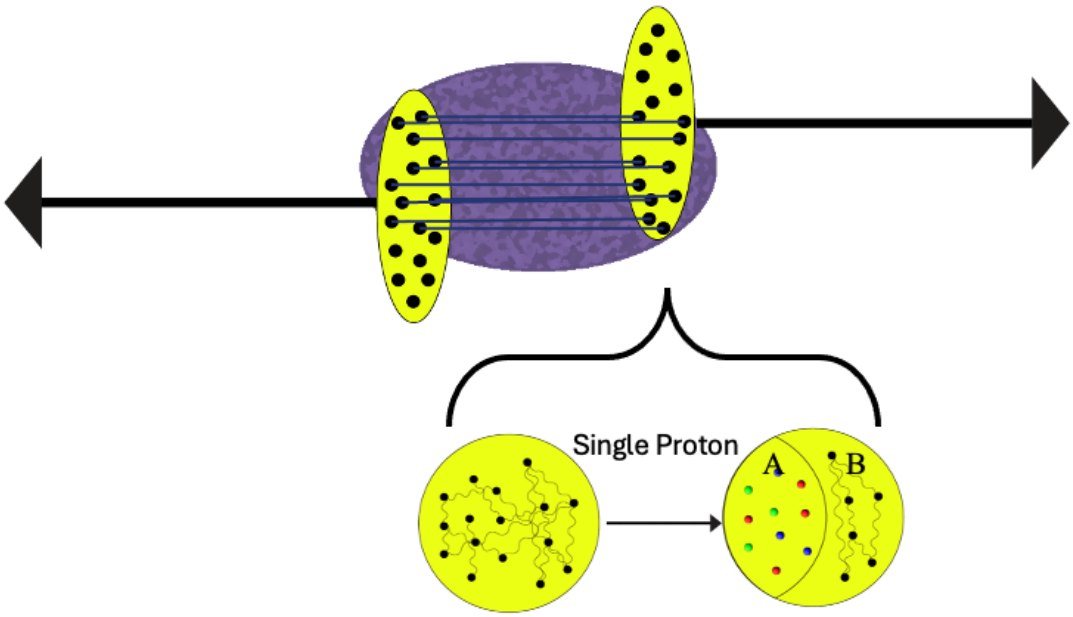

Figure 4.1: Decoherence of the maximally entangled state occurs between the partons probed in the interaction region A and the remainder of the partons in region B.

Information in the initial state is stored in density matrix of the constituent partons. We seek to understand how a quench in the initial state gives rise to the final state distribution. Prior to the collision the proton exists in a pure state, meaning all the constituent partons are

fully entangled. At the moment of collision there is a separation of the wave function between a portion that interacts with the other proton and the portion that is unprobed. This causes decoherence between partons in each region. This decoherence gives rise to a redistribution of information that mimics a thermodynamic distribution. We will neglect entanglement in the longitudinal direction assuming that flux tube merging creates a completely coherent state within the interaction[1]. Therefore, the entropy of the interaction can be simplified to an entropy created at the ends of strings generated by parton-parton interactions.

4.1 Parton Distribution Functions

Parton Distribution Functions (PDFs) are fundamental components in Quantum Chromodynamics (QCD) that describe the probability densities of finding a parton (quark or gluon) within a hadron (such as a proton) carrying a specific fraction of the hadron's momentum. PDFs are crucial for making theoretical predictions for high-energy processes involving hadrons. These distributions are extracted from DIS cross sections as described in chapter 1.

4.1.1 Dependence on the Q^2 Scale

The PDFs are not static but evolve with the energy scale Q^2 of the probing interaction. This evolution is governed by the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) equations, which describe how the PDFs change with Q^2 [68]:

$$\frac{\partial f_i(x,Q^2)}{\partial \ln Q^2} = \sum_j \int_x^1 \frac{dy}{y} P_{ij}(y, \alpha_s(Q^2)) f_j\left(\frac{x}{y}, Q^2\right), \tag{4.1}$$

[1]This approach is supported by the color glass condensate model describe in chapter 1

where $f_i(x, Q^2)$ represents the PDF for parton i, $P_{ij}(y, \alpha_s(Q^2))$ are the splitting functions that describe the probability of a parton j splitting into a parton i with a fraction y of the original parton's momentum, and $\alpha_s(Q^2)$ is the strong coupling constant. The relevant energy scale at which protons are probed in the LHC is called the gluon saturation scale.[2]

4.1.2 Dependence on the Strong Coupling Constant

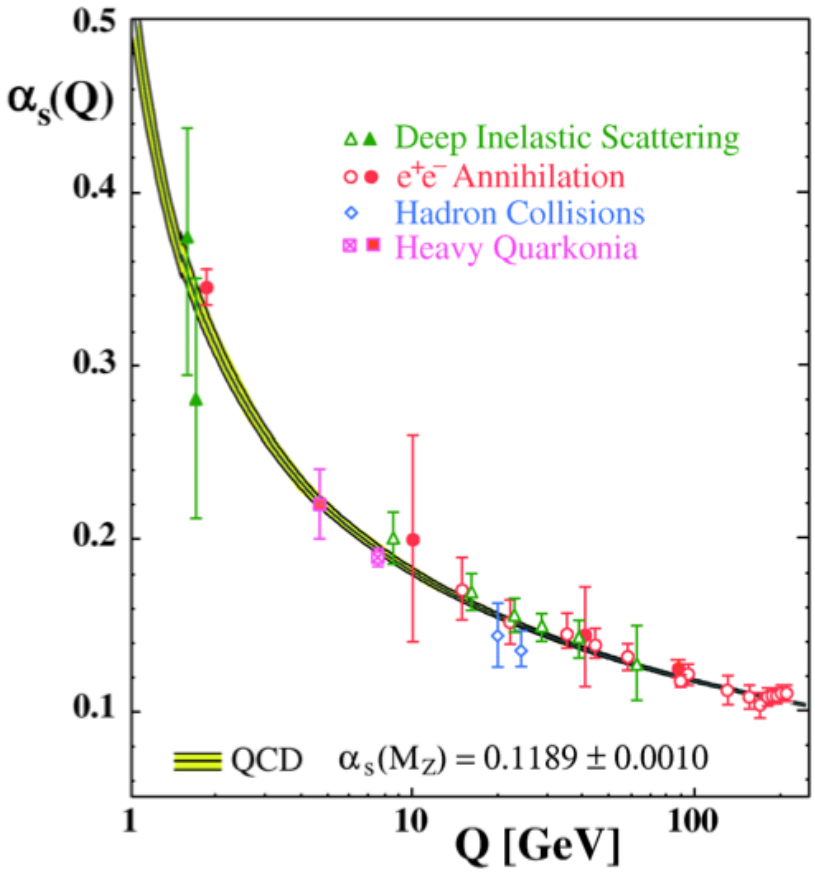

Figure 4.2: Coupling constant as a function of the energy scale Q [25]. Averaging over each of the results gives a baseline coupling constant at the energy scale of the Z boson mass equal to about 0.1189

The strong coupling constant α_s is a crucial parameter in QCD, describing the strength

[2]This scale is defined in chapter 1 and calculated from the average of x as described in chapter 3

of the interaction between quarks and gluons. The value of α_s decreases with increasing Q^2, a phenomenon known as asymptotic freedom. The running of α_s with Q^2 is described by the renormalization group equation [25]:

$$\frac{d\alpha_s(Q^2)}{d\ln Q^2} = -\beta(\alpha_s(Q^2)), \tag{4.2}$$

where the beta function $\beta(\alpha_s)$ can be expanded as:

$$\beta(\alpha_s) = \beta_0\frac{\alpha_s^2}{4\pi} + \beta_1\frac{\alpha_s^3}{(4\pi)^2} + \cdots, \tag{4.3}$$

with $\beta_0 = 11 - \frac{2}{3}n_f$ and $\beta_1 = 102 - \frac{38}{3}n_f$, where n_f is the number of active quark flavors.

The dependence of PDFs on α_s introduces an additional complexity in their determination and evolution, as both the splitting functions $P_{ij}(y, \alpha_s(Q^2))$ and the PDFs themselves are functions of α_s.

Typically PDFs start from a reference scale M_z (M_Z is the mass of the Z boson), which is constant, and the strong coupling constant at any scale Q^2 can be obtained through evolution of the DGLAP equations [69].

The value of $\alpha_s(M_Z)$ has been measured in various experiments, as shown in the above figure there is some disagreement on the magnitude of this constant. Averaging over all predictions gives a value for the reference scale equal to about 0.1189 [25].

4.1.3 Series Expansion: LO, NLO, NNLO

The DGLAP equations and the calculation of cross sections involving PDFs can be expanded in a perturbative series in terms of α_s. These expansions are referred to as Leading Order (LO), Next-to-Leading Order (NLO), and Next-to-Next-to-Leading Order (NNLO), ect. [70].

Leading Order (LO)

At LO, the splitting functions and the cross sections are calculated using the lowest-order Feynman diagrams. This provides a first approximation to the PDFs and their evolution:

$$P_{ij}^{\mathrm{LO}}(y) = P_{ij}^{(0)}(y). \tag{4.4}$$

LO calculations are the simplest and provide a basic understanding of parton dynamics.

Higher Order Expansions

At NLO, the calculations include corrections from one-loop Feynman diagrams. These corrections provide more accurate descriptions of the PDFs and their evolution:

$$P_{ij}^{\mathrm{NLO}}(y) = P_{ij}^{(0)}(y) + \frac{\alpha_s}{2\pi} P_{ij}^{(1)}(y). \tag{4.5}$$

At NNLO, the calculations include corrections from two-loop Feynman diagrams, further improving the accuracy of the PDFs and their evolution:

$$P_{ij}^{\mathrm{NNLO}}(y) = P_{ij}^{(0)}(y) + \frac{\alpha_s}{2\pi} P_{ij}^{(1)}(y) + \left(\frac{\alpha_s}{2\pi}\right)^2 P_{ij}^{(2)}(y). \tag{4.6}$$

4.1.4 Justification for LO at Low Q^2

At low Q^2, the strong coupling constant α_s is large, making the perturbative series expansion less convergent. The LO approximation, which includes only the leading-order terms, provides a simple and robust description without the need for complicated higher-order corrections that may be less accurate in this regime.

Additionally, the gluon density at low Q^2 is often dominated by non-perturbative effects

that are not well-captured by higher-order perturbative QCD calculations. Using the LO approximation allows for a straightforward representation of the gluon distributions while avoiding the complexities and uncertainties associated with higher-order terms.

4.2 Extrapolating PDFs

PDFs reflecting different experimental fits are published by many collaborations. These PDFs are generally published at some defined range in x and starting at a reference scale $\alpha_s(M_Z)$ as described by the previous section. In order to make a comparison to the kinematic range of our final state distributions these initialized PDFs are extrapolated using various interpolation methods and DGLAP evolution calculations.

4.2.1 LHAPDF

The LHAPDF (Les Houches Accord PDF) framework is a software library designed to provide a standardized interface for Parton Distribution Functions (PDFs) used in high-energy physics simulations and calculations [70]. The main aim of the LHAPDF is to facilitate the usage and comparison of various PDF sets from different collaborations, such as MSTW, HERAPDF, CT, and NNPDF. By providing a unified interface to access these PDF sets, LHAPDF greatly simplifies the process of incorporating PDFs into theoretical predictions and experimental analyses.

4.2.2 Features of LHAPDF

The LHAPDF framework offers several key features that make it a valuable tool for high-energy physicists:

- **Unified Interface**: LHAPDF provides a single, consistent interface to access PDF

sets from various groups. This allows users to easily switch between different PDF sets without having to change their code significantly.

- **Interpolation and Extrapolation**: The framework includes interpolation and extrapolation algorithms that enable users to evaluate PDFs at any point in the (x, Q^2) plane within the limits of the PDF sets. This allows for a more accurate determination of parton densities in different kinematic regions, including the low x and low Q^2 regime.

- **Error Estimation**: LHAPDF provides tools for estimating uncertainties associated with the PDFs, which can be propagated to the final results of theoretical predictions or experimental analyses.

- **Compatibility**: The framework is designed to be compatible with various programming languages and platforms, including C++, Fortran, and Python. This makes it accessible to a wide range of users and allows for seamless integration with existing software packages used in high-energy physics.

4.3 Hessian Errors

In high-energy physics, the uncertainties in Parton Distribution Functions (PDFs) can significantly impact the predictions for various observables and cross sections. Estimating these uncertainties is crucial when comparing theoretical predictions to experimental data. One of the widely used methods to estimate PDF uncertainties is the Hessian approach. In this section, we will discuss how the Hessian approach is employed in the LHAPDF framework to predict PDF uncertainties.

The Hessian approach is based on the quadratic expansion of the χ^2 function, which measures the goodness of fit between the experimental data and theoretical predictions.

By expanding the χ^2 function around its minimum, the Hessian approach allows for the determination of uncertainties in the fitted PDF parameters [71].

In the Hessian approach, the PDF uncertainty is represented by a symmetric error matrix, known as the Hessian matrix. The Hessian matrix is defined as the second derivative of the χ^2 function with respect to the PDF parameters:

$$H_{ij} = \frac{\partial^2 \chi^2}{\partial a_i \partial a_j}, \tag{4.7}$$

where a_i and a_j are the PDF parameters. The eigenvectors of the Hessian matrix correspond to the principal directions in the parameter space, while the eigenvalues represent the uncertainties along these directions.

4.3.1 Using the Hessian Approach in LHAPDF

The LHAPDF framework incorporates the Hessian approach to provide a consistent method for estimating PDF uncertainties across different PDF sets [71]. Once the Hessian matrix is obtained from the global fit, a set of eigenvector PDFs is generated, corresponding to the principal directions in the parameter space. Typically, there are $2N$ eigenvector PDFs, where N is the number of independent PDF parameters.

The eigenvector PDFs can be used to estimate the uncertainty in any observable or cross section. For a given observable O, the PDF uncertainty can be calculated as follows:

$$\Delta O = \sqrt{\sum_{i=1}^{N} \left[\frac{O(a_i^+) - O(a_i^-)}{2} \right]^2}, \tag{4.8}$$

where $O(a_i^+)$ and $O(a_i^-)$ are the values of the observable calculated using the i-th pair of eigenvector PDFs, and N is the number of independent PDF parameters.

The LHAPDF framework provides tools and functions to access the eigenvector PDFs

and calculate the uncertainties for various observables and cross sections. By using the Hessian approach, LHAPDF ensures a consistent treatment of PDF uncertainties, allowing for a reliable comparison between theoretical predictions and experimental data.

4.4 Generated PDFs

The following PDFs were extrapolated using the LHAPDF framework. One set extrapolated from published results of the MSHT collaboration using a reference scale of $\alpha_s(M_Z) = 0.13\ GeV^2$ [72]. Two more sets were generated from published results of the NNPDF collaboration, one using the reference scale $\alpha_s(M_Z) = 0.13\ GeV^2$ and another at $\alpha_s(M_Z) = 0.119\ GeV^2$ [73].

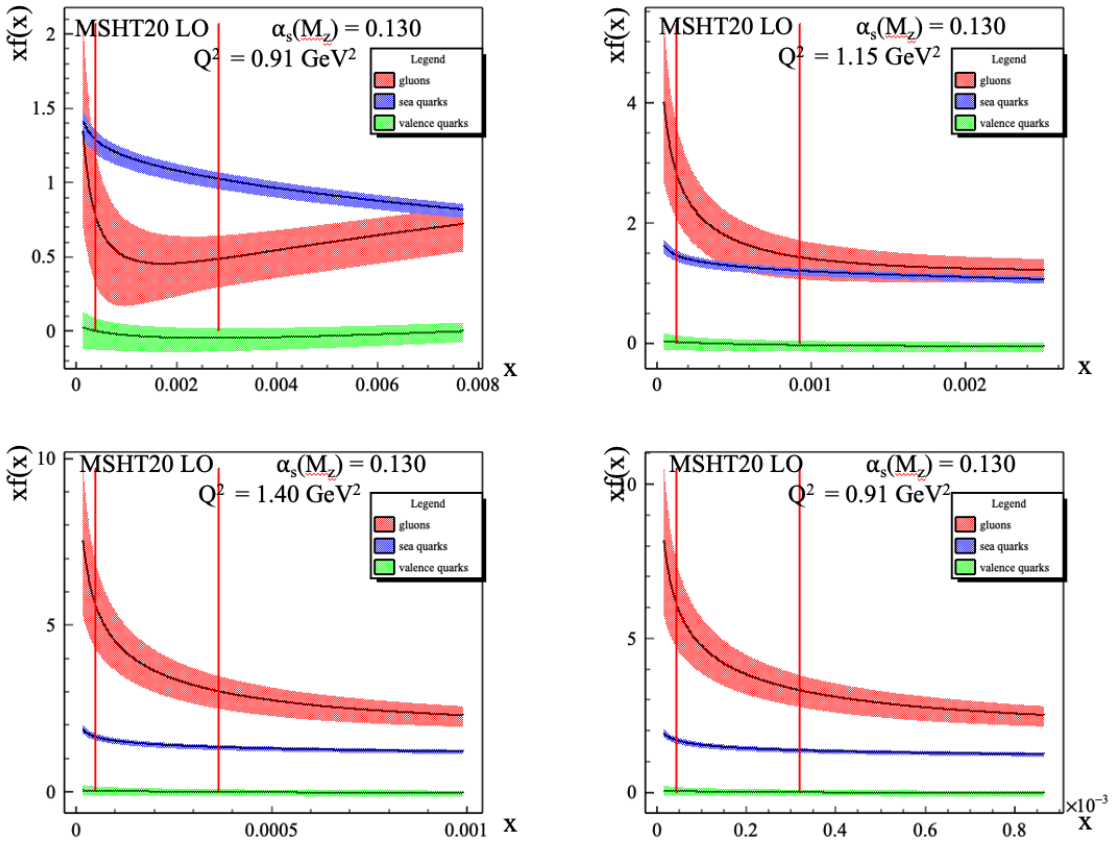

Figure 4.3: PDF set defining the initial state of the proton from the MSHT collaboration at $\alpha_s(M_Z) = 0.13\ GeV^2$. [72]

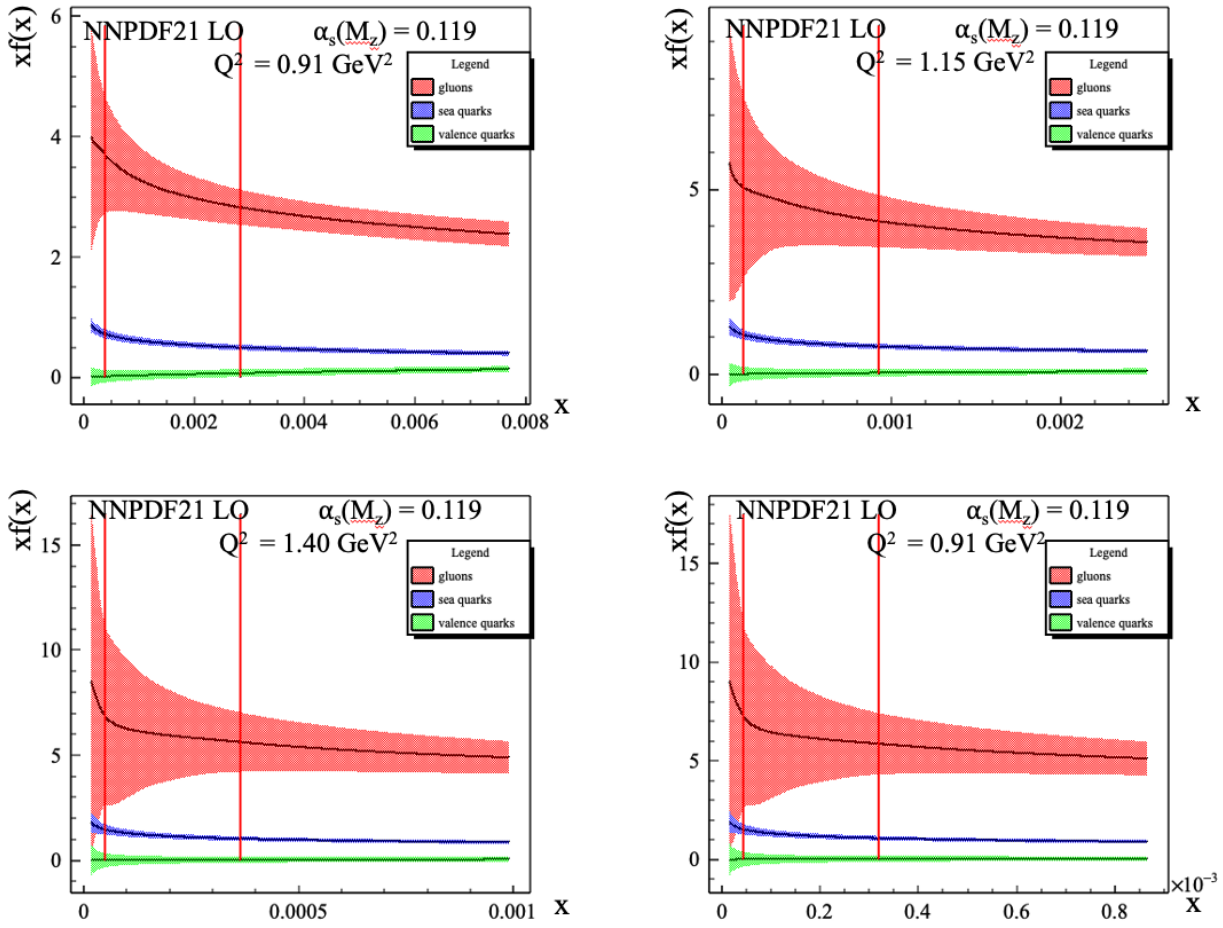

Figure 4.4: PDF set defining the initial state of the proton from the NNPDF collaboration at $\alpha_s(M_Z) = 0.119\ GeV^2$ [73]

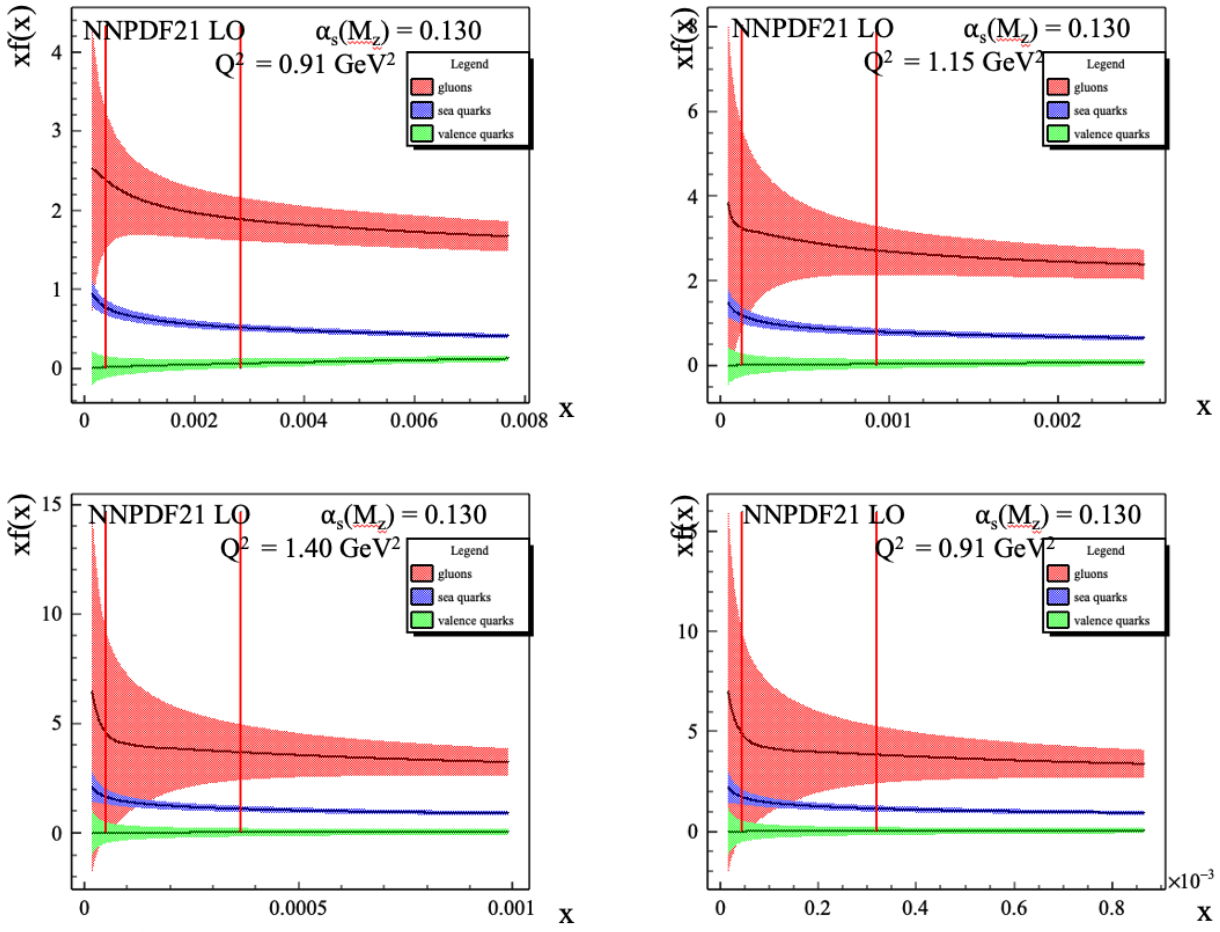

Figure 4.5: PDF set defining the initial state of the proton from the NNPDF collaboration at $\alpha_s(M_Z) = 0.13\ GeV^2$. [73]

In the above figures red bands represent the distribution of gluons, blue bands represent sea quarks, and green bands represent valance quarks. Red vertical lines represent the kinematic range (defined by the range in η in the final state) over which these PDFs are integrated to to get the total number of partons.

4.5 Initial State Entropy

An initial-state entropy can now be calculated based on the distribution of partons defined by the PDFs shown in the previous section. Due to a lack in information contained in the off

diagonal elements of the parton density matrix we will make a calculation of the "entropy of ignorance" defined in chapter 1:

$$S_I = -\mathrm{Tr}\left(\rho_I \ln \rho_I\right). \tag{4.9}$$

Where ρ_I represents a matrix with off-diagonal elements equal to 0. In a system comprised mostly of indistinguishable gluons the degrees of freedom associated with the diagonal elements represent individual partonic states (one for each parton), and the equation reduces to:

$$S_I = ln(N_{partons}) \tag{4.10}$$

Where $N_{partons}$ represents the number of interacting partons in the intial-state calculated from PDFs as follows:

$$N_{\text{partons}} = \int_{x_1}^{x_2} \frac{xg(x)}{x}\, dx + \int_{x_1}^{x_2} \frac{x\Sigma(x)}{x}\, dx \tag{4.11}$$

Where $g(x)$ and $\Sigma(x)$ represent the parton distribution functions $f(x)$ for gluons and quarks respectively, and the range $x_1 - x_2$ represents the kinematic range (set by the pseudorapidity) measured in the final state.

4.5.1 Correction for unmeasured neutral hadrons

Measuring neutral hadrons in high-energy particle collisions presents significant experimental challenges. Unlike charged particles, neutral hadrons do not interact with electromagnetic fields, making their detection and measurement more complex. This section explores the difficulties associated with measuring neutral hadrons and discusses the assumption that approximately one-third of the produced particles are neutral. We will make a correction in

our initial state on the basis of this assumption [53].

Equal Distribution Assumption

In high-energy collisions, charge conservation ensures that the number of positively charged particles is balanced by an equal number of negatively charged particles. Furthermore, the production mechanisms for hadrons are generally symmetric with respect to charge. In processes such as hadronization, where quark-antiquark pairs are produced, the probabilities of forming positive, negative, and neutral hadrons are expected to be similar. The assumption can be mathematically expressed as:

$$N_{\text{total}} = N_{\text{positive}} + N_{\text{negative}} + N_{\text{neutral}}, \tag{4.12}$$

where N_{total} is the total number of produced particles, N_{positive} is the number of positively charged particles, N_{negative} is the number of negatively charged particles, and N_{neutral} is the number of neutral particles. Assuming equal distribution, we have:

$$N_{\text{positive}} \approx N_{\text{negative}} \approx N_{\text{neutral}} \approx \frac{N_{\text{total}}}{3}. \tag{4.13}$$

Experimental data from high-energy collisions, such as those conducted at the Large Hadron Collider (LHC) and other facilities, support the assumption of an approximately equal distribution of positive, negative, and neutral hadrons. While direct measurements of neutral hadrons are challenging, indirect evidence from charged particle measurements and overall particle multiplicities supports this assumption.

Implementation

Because we feel confident in assuming 1/3 of the produced hadrons are neutral, and these are not considered in the final state charged particle distributions we will make a correction to account for this. We will make this correction by assuming one-third of the partons in the initial state are involved in the production of neutral particles and we will remove that one-third. The calculation of the initial state entropy then becomes:

$$S = ln(2/3(N_{gluons} + N_{quarks})) = ln(N_{gluons} + N_{quarks}) + ln(2/3) \tag{4.14}$$

Chapter 5

Final Results and Conclusions

This chapter will discuss the final results of the multiplicity analysis for p-p collisions at 13 TeV. It will also compare the information spread and entanglement entropy between the final and initial states. A final interpretation of the results will be made, including conclusions drawn about entropy generated from different partonic sources and how the strength of the strong coupling constant affects the initial entropy calculation. Finally we will address open questions and the need for future studies.

5.1 Multiplicity at 13 TeV

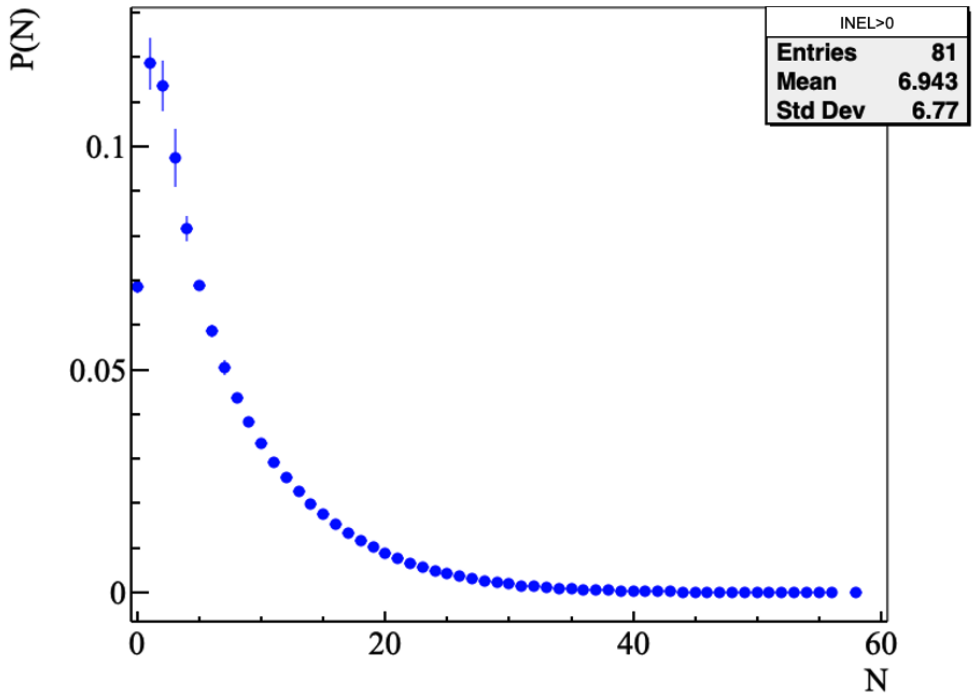

Figure 5.1: Final result for the multiplicity distribution for the INEL>0 event class at $-0.5 < \eta < 0.5$

Analysing ALICE data at 13 TeV for the INEL>0 event class yields the distribution shown above. This distribution is corrected for detector inefficiencies and influences. Errors reflect both statistical and systematic sources of uncertainty, though they are dominated by systematic errors.

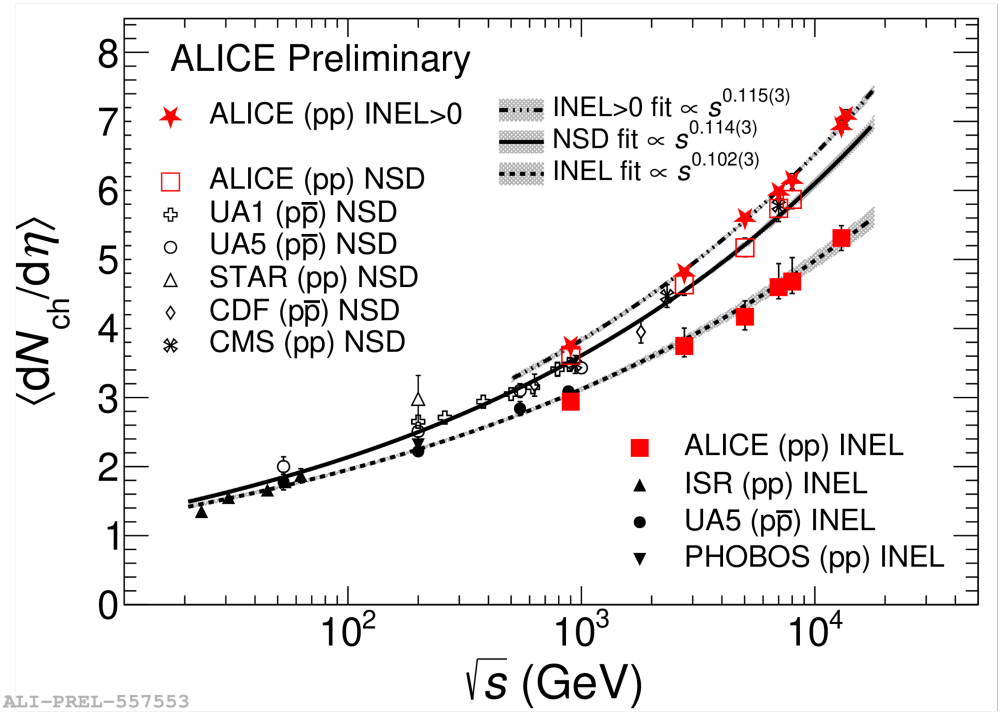

Figure 5.2: Average multiplicity at mid rapidity for all energies follows a power law fit.

In has been shown that average multiplicity at mid-rapidity follows a power law trend. In the context of this work the increase can be attributed to the increase in the number of participating partons. As the energy in the collision increases we probe deeper into the protons wave function where a steep rise the the number of gluons is expected. Results from this analysis are consistent with the expected average multiplicity derived from a power-law fitting of the data.

5.2 Information Spread in the Initial and Final States

Using the 1+1 toy model of non-linear QCD evolution of the BK equation[1], Kharzeev and Levin demonstrated that one can construct a generating function that captures the non-linear interactions leading to dipole formation, in an entangled system, at different rapidities [45].

[1]This model was described in chapter 1

Using this generating function and assuming an NBD distribution in the final state one can calculate the expected moments of the final state distribution.

$$C_2 = 2 - \frac{1}{\bar{n}} \tag{5.1}$$

$$C_3 = \frac{6(\bar{n}-1)\bar{n}+1}{\bar{n}^2} \tag{5.2}$$

$$C_4 = \frac{(12\bar{n}(\bar{n}-1)+1)(2\bar{n}-1)}{\bar{n}^3} \tag{5.3}$$

$$C_5 = \frac{(\bar{n}-1)(120\bar{n}^2(\bar{n}-1)+30\bar{n})+1}{\bar{n}^4} \tag{5.4}$$

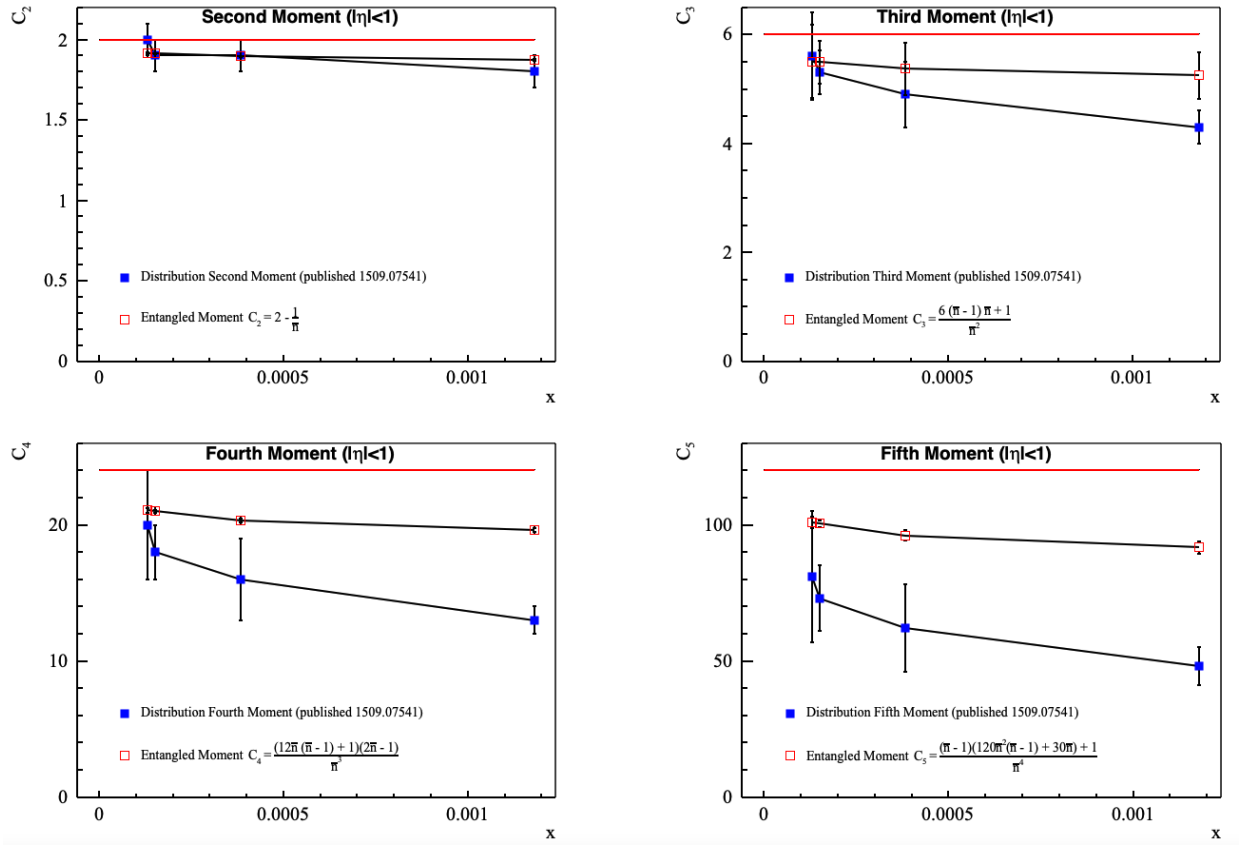

Figure 5.3: Comparison of theoretical moments calculated from a 1+1 toy model of particle production using color dipoles as described in chapter 1 [45], to final-state moments of the multiplicity distribution.

Statistical moments describe the spread of a distribution. In information theory we can think of these moments as describing the spread of information in the system. Higher order moments are more sensitive to this information spread. Comparing the moments from AdS-CFT calculations to the experimentally measured moments, we observe relatively strong agreement in the information spread, particularly at low-x.

Agreement seen here offers an indication that there is no significant increase in information spread during the systems evolution. This points to the possibility of entanglement surviving the systems evolution, and decoherence being the driving mechanism behind matter generation.

5.3 Comparing the Final and Initial State Entropy

As described in the previous chapters of this work entropy provides a metric for understanding the degree of information loss in a system. In the context of p-p collisions this information loss happens due to decoherence in the initial entangled state of partons. Entropy also relates to thermodynamic properties such as temperature, which is a macroscopic quantity. We seek to resolve a correspondence between the purely quantum system found in the initial state, and a thermal-like system in the final state displaying macroscopic collective behavior.

5.3.1 Quark Contribution to Entropy

In the initial paper defining the measurement of entropy in p-p collisions it was believed that quark contributions to the initial-state entropy were negligible [51]. However, we observe a significant rise in the initial entropy when taking into account the contributions from both

gluons, sea quarks, and valence quarks as shown in the figure below.[2]

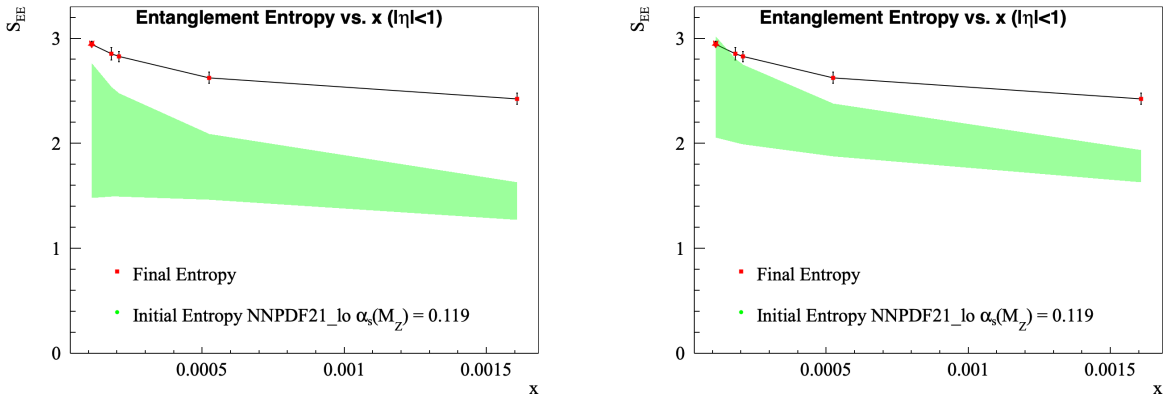

Figure 5.4: Comparison of final-state entropy to initial state entropy (a) without quark contribution (b) with all quark contributions.

5.3.2 Correcting for unmeasured degrees of freedom

Duan et. all made a theoretical calculation relating the entropy of ignorance to the true entanglement entropy [52]. In the context of the Color Glass Condensate (CGC) framework.

The calculated ratio S_I/S_E, which varies with the transverse momentum q, as shown in the figure below, approaches unity at large q. This disparity between the two forms of entropy arises from the experiments inability to measure all correlated degrees of freedom.

[2]Valence quark contributions are near 0, but the sea quarks contribute significantly to the entropy.

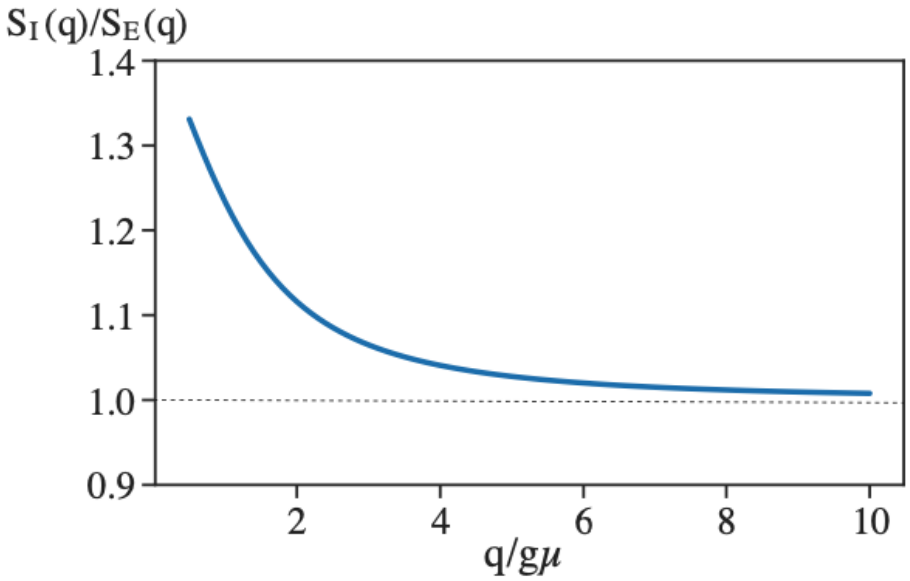

Figure 5.5: Ratio between the entropy of ignorance and entanglement entropy in the initial state. [52]

Because we believe there is no entropy increase during the systems evolution, the final state entropy would be equal to the entanglement entropy reflective of the full density matrix in the initial state. However, our calculation of entropy only considers the diagonal elements of the density matrix[3], as defined by the CGC model, and therefore not reflective of the true entanglement entropy. A correction on the initial-state is made to account for these unmeasured degrees of freedom thus increasing the final state entropy. This correction becomes negligible at high Q^2 and therefore not relevant in past analyses using DIS data which is binned in $Q^2 > 5GeV^2$.

[3]Off diagonal elements of the density matrix representing correlations between unmeasured degrees of freedom in our experiment.

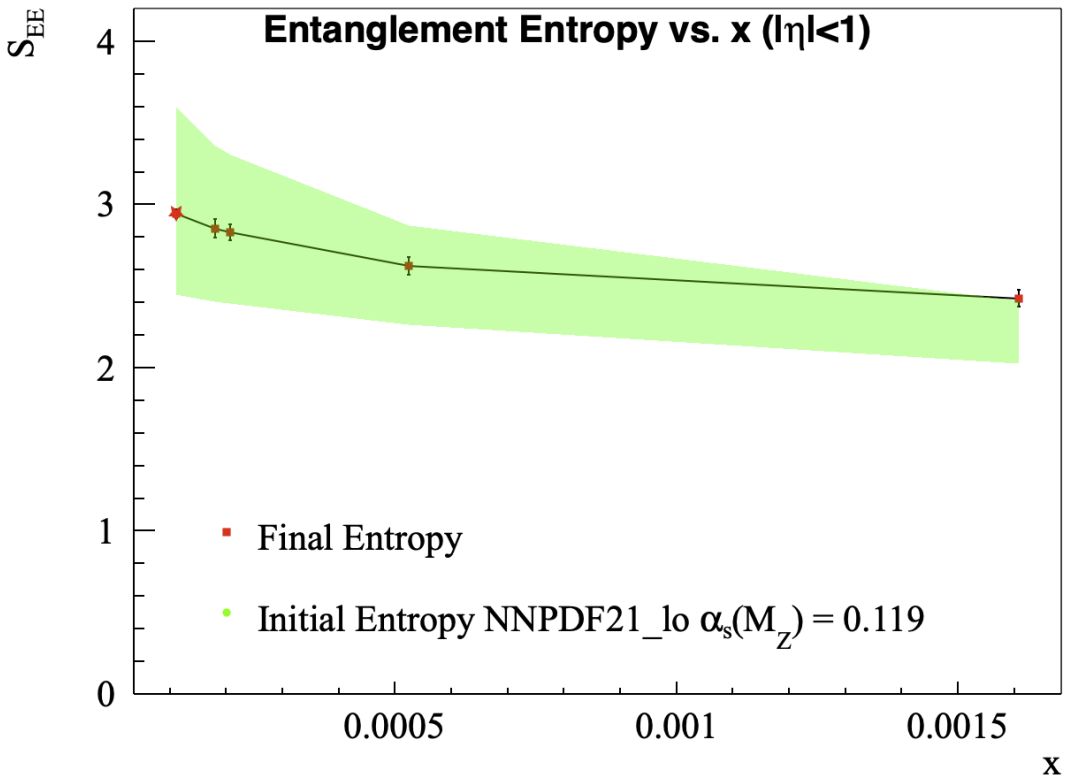

Figure 5.6: Final state entropy including all corrections compared to initial-state entropy with all partonic contributions.

After making the correction we see strong agreement within errors between the initial and final state entropy. Again we observe a stronger correspondence at low-x. This provides further evidence that the entanglement is preserved giving rise to the final state characterized by macroscopic behavior.

5.4 Verification of a Low $\alpha_s(M_Z)$

As described in the previous chapter there exists some discrepancy on the value of the reference scale $\alpha_s(M_Z)$ used in the calculation of PDFs and other quantities. We compared the final-state entropy to the initial-state entropy with three different PDF sets using different values of the reference scale $\alpha_s(M_Z)$.

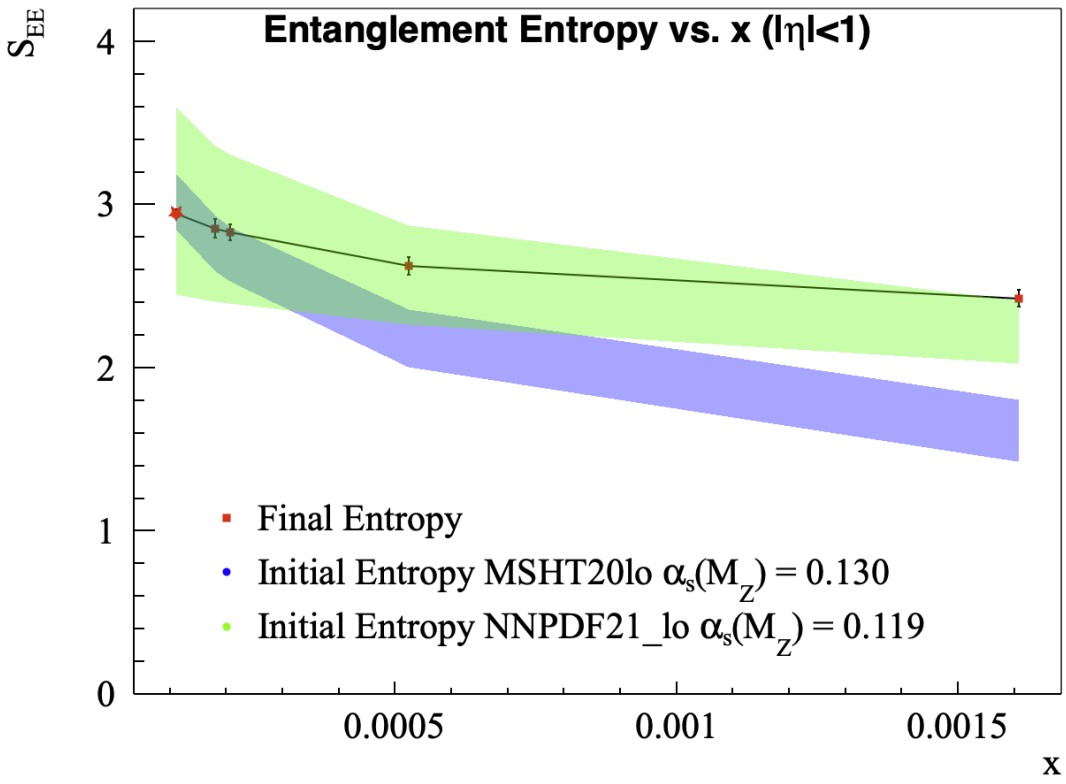

Figure 5.7: Final state entropy including all corrections in red, and the initial state entropy shown by colored bands for 3 models at different values of $\alpha_s(M_Z)$.

Results show stronger agreement between the entropy for $\alpha_s(M_Z) = 0.119\ GeV^2$ when compared to a higher value $\alpha_s(M_Z) = 0.13\ GeV^2$. This value is closer to the global average which is reported to be approx $0.1189\ GeV^2$ [25]. Our results offer further evidence of a $\alpha_s(M_Z)$ value near the average.

5.5 Future Studies

This study has offered some preliminary evidence of entanglement surviving the evolution of a hadronic collision. However, it has yet to be seen if this correspondence exists in larger systems which irrefutably generate a QGP. Further studies in these larger systems are necessary to elucidate the idea of entanglement surviving this phase of strongly interacting

matter.

Additionally, studies of entanglement entropy should be done more differentially in x to explore the full Page curve. Studies should also be done using particle identification to understand how entanglement plays a role in the species of generated matter and how quarks and gluons individually contribute each cross section.

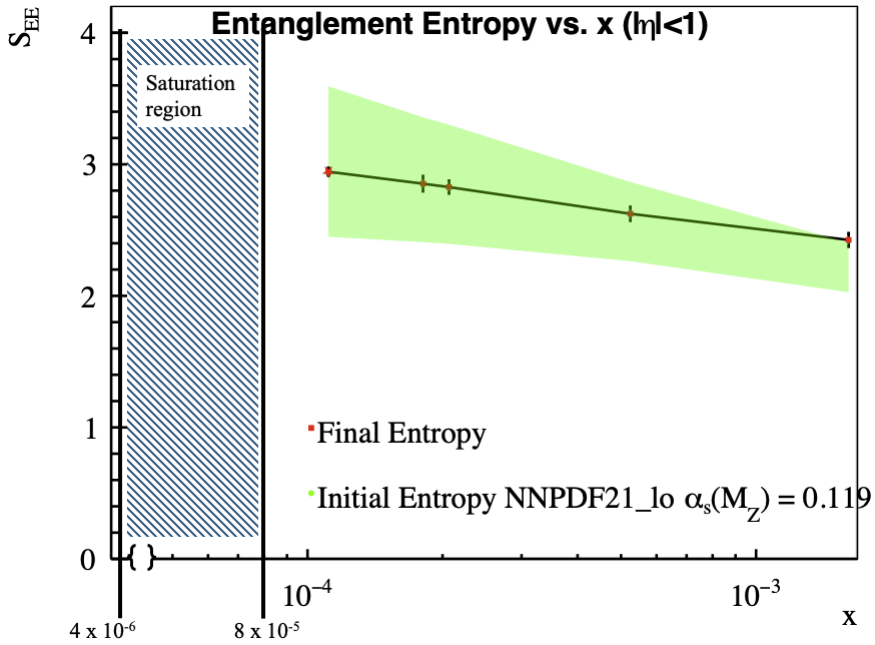

Figure 5.8: Entropy as a function of x on a log scale, showing the region where saturation should be observed in future studies.

Finally, in the theoretical framework of a CGC there should exist gluons saturation at low-x this has yet to be seen and should be evident in the entanglement entropy. Studies should be done at lower values of x, a kinematic regime which will be further explored by the Electron Ion Collider (EIC) [74] in the coming years.

5.6 Conclusion

It is understood that a system involving colliding hadrons at high energies behaves like a thermodynamic process. In order to model the hydrodynamic-like behavior that is observed within these collisions very early thermalization is required. This early thermalization can be explained using quantum entanglement. In the final state a direct connection can be drawn between the entropy used to define macroscopic properties (like temperature) and the entanglement entropy generated from a quantum system.

In order to prove these hypotheses, and verify previous studies in simpler systems (e-p) we look for signatures of entanglement in hadronic collisions. The smallest system that displays collective behavior, indicative of QGP formation, are high energy p-p collisions. In an effort to provide experimental evidence of entanglement both surviving the system evolution and driving the final state particle yields we looked for signatures of entanglement within the distribution of final-state hadrons produced in p-p collisions at the LHC.

Comparing the entanglement entropy and information spread in the initial state (defined by model predictions fitted to data (PDFs)) and the final state, we find a strong correspondence. This correspondence indicates that entanglement does in fact survive the system evolution, and has a considerable effect on matter generation.

Experimental evidence has shown that high energy hadronic collisions behave like a hydrodynamic system in thermal equilibrium. Furthermore, the entropy calculated in the final state is proportional to the thermodynamic entropy from which one can derive a temperature in a thermalized system. Agreement between the initial-state entanglement entropy, derived from fundamental quantum principles and QCD, and the final-state entropy offers a powerful example of macroscopic behavior as emergent behavior from a quantum system.

Finally, high energy collisions are often used in understanding the Big Bang and early evolution of the universe. In the context of this study, and its implications in describing

the early universe, some interesting philosophical conclusions might be drawn. Prior to the initial bang that began the universe as we know it, it is believed all the energy of the universe was contained in a point like object. This singularity is understood to be perfectly symmetric and coherent. Much like our proton, the universe represented a pure quantum state with all possible configurations of the universe contained in its wavefunction. While it is not understood what caused the Big Bang, if it was a similar process to that of hadronic collisions, one might conclude that entanglement in that initial state drove the expansion and the distribution of matter in the very early universe.[4] Such behavior would suggest a universe set in motion by the inexorable pull of fate.

[4]A continued thermodynamic evolution of the universe over time would undoubtedly destroy evidence of coherence in the initial state, making such a conjecture impossible to prove in our current understanding of physics. Physics without some framework for measurement and experimental proof is not physics, making this a purely metaphysical and philosophical statement.

Bibliography

[1] A. Einstein, B. Podolsky, and N. Rosen, "Can quantum-mechanical description of physical reality be considered complete?," *Physical Review*, vol. 47, pp. 777–780, 1935.

[2] E. Schrödinger, "The present status of quantum mechanics," *Naturwissenschaften*, vol. 23, pp. 807–812, 823–828, 844–849, 1935.

[3] G. Adesso, "The social aspects of quantum entanglement," 2007.

[4] M. Bauer, "The stern-gerlach experiment, translation of: "der experimentelle nachweis der richtungsquantelung im magnetfeld"," 2023.

[5] A. Aspect, P. Grangier, and G. Roger, "Experimental realization of einstein-podolsky-rosen-bohm gedankenexperiment: A new violation of bell's inequalities," *Physical Review Letters*, vol. 49, no. 2, pp. 91–94, 1982.

[6] W. H. Zurek, "Decoherence, einselection, and the quantum origins of the classical," *Reviews of Modern Physics*, vol. 75, p. 715–775, May 2003.

[7] R. Landauer, "Irreversibility and heat generation in the computing process," *IBM Journal of Research and Development*, vol. 5, no. 3, pp. 183–191, 1961.

[8] A. Peres, "Separability criterion for density matrices," *Physical Review Letters*, vol. 77, p. 1413–1415, Aug. 1996.

[9] X.-D. Dai, Z. Wang, H.-R. Wang, and Z. Wang, "Steady-state topological order," *arXiv preprint arXiv:2310.17612*, 2023.

[10] J. M. Maldacena, "The large n limit of superconformal field theories and supergravity," *International Journal of Theoretical Physics*, vol. 38, pp. 1113–1133, 1999. Adv. Theor. Math. Phys. 2 (1998) 231-252.

[11] "Standard model of elementary particles," *Wikipedia, The Free Encyclopedia*, 2019.

[12] P. Langacker, "Structure of the standard model," 2003.

[13] A. Pich, "The standard model of electroweak interactions," 2012.

[14] H. Yukawa, "On the interaction of elementary particles. i.," *Proceedings of the Physico-Mathematical Society of Japan. 3rd Series*, vol. 17, pp. 48–57, 1935. Read Nov. 17, 1934.

[15] A. Collaboration, "Observation of a new particle in the search for the standard model higgs boson with the atlas detector at the lhc," *Physics Letters B*, vol. 716, p. 1–29, Sept. 2012.

[16] "Precision electroweak measurements on the z resonance," *Physics Reports*, vol. 427, p. 257–454, May 2006.

[17] Y. F. e. a. The Super-Kamiokande Collaboration, "Evidence for oscillation of atmospheric neutrinos," *Physical Review Letters*, vol. 81, p. 1562–1567, Aug. 1998.

[18] E. Rutherford, "Collision of alpha particles with light atoms. iv. an anomalous effect in nitrogen," *Philosophical Magazine*, vol. 37, pp. 581–587, 1919.

[19] R. P. Feynman, "Very high-energy collisions of hadrons," *Physical Review Letters*, vol. 23, pp. 1415–1417, 1969.

[20] J. I. Friedman, H. W. Kendall, and R. E. Taylor, "Observed behavior of highly inelastic electron-proton scattering," *Physical Review Letters*, vol. 23, pp. 930–934, 1969.

[21] M. Breidenbach, J. I. Friedman, H. W. Kendall, R. E. Taylor, P. F. Schultz, E. D. Bloom, D. H. Coward, H. C. DeStaebler, J. D. Bjorken, and L. W. Mo, "Deep inelastic scattering of electrons on protons and the observation of scaling," *Physical Review Letters*, vol. 23, pp. 935–939, 1969.

[22] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, "Parton distributions for the lhc," *The European Physical Journal C*, vol. 63, p. 189–285, July 2009.

[23] D. J. Gross and F. Wilczek, "Ultraviolet behavior of non-abelian gauge theories," *Physical Review Letters*, vol. 30, pp. 1343–1346, 1973.

[24] H. D. Politzer, "Reliable perturbative results for strong interactions?," *Physical Review Letters*, vol. 30, pp. 1346–1349, 1973.

[25] S. Bethke, "Experimental tests of asymptotic freedom," *Progress in Particle and Nuclear Physics*, vol. 58, p. 351–386, Apr. 2007.

[26] F. Knechtli, "Lattice quantum chromodynamics," 2017.

[27] S. Collaboration, "Studying the phase diagram of qcd matter at rhic." STAR White Paper, June 2014. 01 June 2014.

[28] S. Borsanyi, Z. Fodor, J. N. Guenther, R. Kara, S. D. Katz, P. Parotto, A. Pasztor, C. Ratti, and K. K. Szabó, "Qcd crossover at finite chemical potential from lattice simulations," *Physical Review Letters*, vol. 125, July 2020.

[29] A. Einstein, "Ist die trägheit eines körpers von seinem energieinhalt abhängig?," *Annalen der Physik*, vol. 323, no. 13, pp. 639–641, 1905.

[30] H. Petersen, *Measurements of differential cross sections for the production of top quark pairs in pp collisions at s = 13 TeV with the CMS experiment*. PhD thesis, CERN, 2022.

[31] B. Foster, R. Thorne, and M. Vincter, "Structure functions," *Particle Data Group*, September 2017.

[32] Z. C. The H1, "Combined measurement and qcd analysis of the inclusive e ± p scattering cross sections at hera," *Journal of High Energy Physics*, vol. 2010, Jan. 2010.

[33] B. Müller and A. Schäfer, "Quark-hadron transition and entanglement," 2022.

[34] D.-M. Zhou, Y.-L. Yan, B.-G. Dong, X.-M. Li, D.-J. Wang, X. Cai, and B.-H. Sa, "Charged particle elliptic flow in collisions at lhc energies in a parton and hadron cascade model paciae," *Nuclear Physics A*, vol. 860, p. 68–75, June 2011.

[35] U. W. Heinz and P. F. Kolb, "Two rhic puzzles: Early thermalization and the hbt problem," 2002.

[36] A. Collaboration, "Emergence of long-range angular correlations in low-multiplicity proton-proton collisions," 2023.

[37] A. Collaboration, "Enhanced production of multi-strange hadrons in high-multiplicity proton–proton collisions," *Nature Physics*, vol. 13, p. 535–539, Apr. 2017.

[38] L. Turko, "Statistical models of hadron production—simple models for complicated processes," *Physics of Particles and Nuclei*, vol. 39, p. 1095–1101, Nov. 2008.

[39] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equation of state calculations by fast computing machines," *The Journal of Chemical Physics*, vol. 21, no. 6, pp. 1087–1092, 1953.

[40] C. Bierlich, S. Chakraborty, N. Desai, L. Gellersen, I. Helenius, P. Ilten, L. Lönnblad, S. Mrenna, S. Prestel, C. Preuss, T. Sjöstrand, P. Skands, M. Utheim, and R. Verheyen, "A comprehensive guide to the physics and usage of pythia 8.3," 2022.

[41] P. Shukla, "Glauber model for heavy ion collisions from low energies to high energies," 2001.

[42] A. Andronic, P. Braun-Munzinger, K. Redlich, and J. Stachel, "Hadron yields in central nucleus-nucleus collisions, the statistical hadronization model and the qcd phase diagram," 2021.

[43] F. Gelis, E. Iancu, J. Jalilian-Marian, and R. Venugopalan, "The color glass condensate," *Annual Review of Nuclear and Particle Science*, vol. 60, p. 463–489, Nov. 2010.

[44] E. G. de Oliveira, "Balitsky-kovchegov evolution equation." http://www.if.ufrgs.br/gfpae, 2024. High Energy Phenomenology Group, Instituto de Física, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil.

[45] D. E. Kharzeev and E. M. Levin, "Deep inelastic scattering as a probe of entanglement," *Physical Review D*, vol. 95, June 2017.

[46] F. Becattini, P. Castorina, A. Milov, and H. Satz, "A comparative analysis of statistical hadron production," *The European Physical Journal C*, vol. 66, p. 377–386, Feb. 2010.

[47] A. M. Kaufman, M. E. Tai, A. Lukin, M. Rispoli, R. Schittko, P. M. Preiss, and M. Greiner, "Quantum thermalization through entanglement in an isolated many-body system," *Science*, vol. 353, p. 794–800, Aug. 2016.

[48] J. Kong, R. Jiménez-Martínez, C. Troullinou, V. G. Lucivero, G. Tóth, and M. W. Mitchell, "Measurement-induced, spatially-extended entanglement in a hot, strongly-interacting atomic system," *Nature Communications*, vol. 11, p. 2415, 2020.

[49] F. Reif, *Fundamentals of Statistical and Thermal Physics*. New York: McGraw-Hill, 1965.

[50] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, 1948.

[51] Z. Tu, D. E. Kharzeev, and T. Ullrich, "Einstein-podolsky-rosen paradox and quantum entanglement at subnucleonic scales," *Physical Review Letters*, vol. 124, Feb. 2020.

[52] H. Duan, C. Akkaya, A. Kovner, and V. V. Skokov, "Entanglement, partial set of measurements, and diagonality of the density matrix in the parton model," *Physical Review D*, vol. 101, Feb. 2020.

[53] M. Hentschinski and K. Kutak, "Evidence for the maximally entangled low x proton in deep inelastic scattering from h1 data," *The European Physical Journal C*, vol. 82, Feb. 2022.

[54] A. Collaboration, "The atlas experiment at the cern large hadron collider," *Journal of Instrumentation*, vol. 3, p. S08003, 2008.

[55] A. Safonov, "Cms experiment at the lhc: Commissioning and early physics," 2010.

[56] L. Collaboration, "The lhcb detector at the lhc," *Journal of Instrumentation*, vol. 3, p. S08005, 2008.

[57] A. Collaboration, "The alice experiment at the cern lhc," *Journal of Instrumentation*, vol. 3, p. S08002, 2008.

[58] ALICE Collaboration, "ALICE Institutes." https://alice-collaboration.web.cern.ch/collaboration/alice_institute. Accessed: 2024-06-15.

[59] ALICE Collaboration, "Fast Interaction Trigger (FIT)." https://alice-collaboration.web.cern.ch/menu_proj_items/FIT. Accessed: 2024-06-15.

[60] ALICE Collaboration, "Inner Tracking System (ITS)." https://alice-collaboration.web.cern.ch/ALICE_ITS. Accessed: 2024-06-15.

[61] ALICE Collaboration, "Time Projection Chamber (TPC)." https://alice-collaboration.web.cern.ch/node/34960. Accessed: 2024-06-15.

[62] A. Collaboration, "Charged-particle multiplicities in proton–proton collisions at $\sqrt{s}$ = 0.9 to 8 tev," *The European Physical Journal C*, vol. 77, Jan. 2017.

[63] S. Kullback and R. A. Leibler, "On information and sufficiency," *The Annals of Mathematical Statistics*, vol. 22, no. 1, pp. 79–86, 1951.

[64] J. L. Albacete, A. Dumitru, H. Fujii, and Y. Nara, "Cgc predictions for p+pb collisions at the lhc," *Nuclear Physics A*, vol. 897, p. 1–27, Jan. 2013.

[65] A. Sytov, L. Bandiera, K. Cho, G. A. P. Cirrone, S. Guatelli, V. Haurylavets, S. Hwang, V. Ivanchenko, L. Pandola, A. Rosenfeld, and V. Tikhomirov, "Geant4 simulation model of electromagnetic processes in oriented crystals for accelerator physics," *Journal of the Korean Physical Society*, vol. 83, p. 132–139, June 2023.

[66] G. D'Agostini, "A multidimensional unfolding method based on bayes' theorem," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 362, pp. 487–498, 1995.

[67] T. Pierog, I. Karpenko, J. M. Katzy, E. Yatsenko, and K. Werner, "Epos lhc: Test of collective hadronization with data measured at the cern large hadron collider," *Physical Review C*, vol. 92, Sept. 2015.

[68] M. Markovych and A. Tandogan, "Analytic evolution of dglap equations," 2023.

[69] P. Skands, "Introduction to qcd," in *Searching for New Physics at Small and Large Scales*, WORLD SCIENTIFIC, Sept. 2013.

[70] A. Buckley, J. Ferrando, S. Lloyd, K. Nordström, B. Page, M. Rüfenacht, M. Schönherr, and G. Watt, "Lhapdf6: parton density access in the lhc precision era," *The European Physical Journal C*, vol. 75, Mar. 2015.

[71] J. Pumplin, D. Stump, R. Brock, D. Casey, J. Huston, J. Kalk, H. L. Lai, and W. K. Tung, "Uncertainties of predictions from parton distribution functions. ii. the hessian method," *Physical Review D*, vol. 65, Dec. 2001.

[72] S. Bailey, T. Cridge, L. A. Harland-Lang, A. D. Martin, and R. Thorne, "The european physical journal c," *The European Physical Journal C*, vol. 81, p. 341, 2021.

[73] T. N. C. R. D. Ball, V. Bertone, F. Cerutti, L. D. Debbio, S. Forte, A. Guffanti, J. I.Latorre, J. Rojo, and M. Ubiali, "Nucl. phys. b," *Nucl. Phys. B*, vol. 855, no. 2, p. 153, 2011.

[74] R. A. K. et. al., "Snowmass 2021 white paper: Electron ion collider for high energy physics," 2022.