INITIAL STATE STRUCTURES AND FINAL STATE CORRELATIONS IN HEAVY ION COLLISIONS

Alba Soto Ontoso

Universidad de Granada & Goethe-Universität

April 2018

Advisors:
Prof. Javier López Albacete, Prof. Hannah Petersen

Departamento de Física Teórica y del Cosmos & CAFPE

Programa de Doctorado en Física y Ciencias del Espacio

Editor: Universidad de Granada. Tesis Doctorales
Autor: Alba Soto Ontoso
ISBN: 978-84-9163-929-9
URI: http://hdl.handle.net/10481/52433

La doctoranda/*The doctoral candidate* **Alba Soto Ontoso** y los directores de la tesis/and the thesis supervisors **Javier López Albacete, Hannah Petersen**:

Garantizamos, al firmar esta tesis doctoral, que el trabajo ha sido realizado por la doctoranda bajo la dirección de los directores de la tesis y hasta donde nuestro conocimiento alcanza, en la realización del trabajo, se han respetado los derechos de otros autores a ser citados, cuando se han utilizado sus resultados o publicaciones.

Guarantee, by signing this doctoral thesis, that the work has been done by the doctoral candidate under the direction of the thesis supervisor and, as far as our knowledge reaches, in the performance of the work, the rights of other authors to be cited (when their results or publications have been used) have been respected.

En Frankfurt am Main, a 21 de Marzo de 2018.

Directores de la tesis Doctoranda

Fdo: Javier López Albacete, Hannah Petersen Alba Soto Ontoso

D. Javier López Albacete y D^{a}. Hannah Petersen

CERTIFICAN: que la presente tesis doctoral, INITIAL STATE STRUCTURES AND FINAL STATE CORRELATIONS IN HEAVY ION COLLISIONS, ha sido realizada por D^{a}. Alba Soto Ontoso bajo su dirección en el Dpto. de Física Teórica y del Cosmos de la Universidad de Granada y el Frankfurt Institute for Advance Studies (Alemania) en régimen de cotutela.

En Frankfurt am Main, a 21 de marzo de 2018.

Directores de la tesis

Fdo: Javier López Albacete, Hannah Petersen

El aire de la calle;
Los Delinquentes.

Acknowledgements

"You have two months to write your PhD thesis". The most natural reaction to that asseveration can include heart rate acceleration, cold sweat, trembling legs, all spiced up with some notes of panic. In the end, after swallowing and grinding my teeth, the mission has been accomplished. From the beginning, I was completely sure that this was going to be the section I would most enjoy writing. The time has come to thank all those who have shared with me *el aire de la calle* during the realization of this thesis and since much earlier.

"You are going to have two supervisors". Once again, the most natural reaction...Wait. Nothing could be further from the truth. It has been a real pleasure to work these years together with Javier and Hannah. The two of them have offered me their undeniable support to overcome the different kinds of setbacks that have arisen during the realization of this thesis. The following pages are a clear example of all the physics that they have taught me, so here I would like to focus on other things I have learnt and admire from them. Javier is a natural born fighter who has shown me the benefits of temperance and patience both in science and in life. Discussing with him through a screen or in front of a blackboard has always been tremendously enriching. Hannah bet on this project and warmly hosted me in her group almost without knowing me. Her leadership, management and organizational skills are worthy of admiration. Definitely, the world would be a better place with more women like her in positions of responsibility. Thank you both for enroling this adventure.

This thesis has been carried out between two sites: Granada and Frankfurt. Maybe more if one takes into account planes, trains and hotel rooms where parts of this work was performed. First, following the chronological order of my trips, I would like to express my deepest gratitude to all the members of the Departamento de Física Teórica y del Cosmos of the Universidad de Granada. Besides the very peaceful working environment, I really enjoyed the padel matches at noon under a scorching sun, the attempts to stop a zombie apocalypse and the political discussions at lunch time. Special thanks to José Ignacio because he sparked my interest in particle physics with his brilliant lectures and patiently guided my first steps in this research business. Antonio's work attitude is an example to follow for any junior scientist. It has been an enormous delight to take part in the gastronomic activities of the experimental family. Thanks for not taking too seriously our business ideas amongst other things, otherwise this thesis may have never been finalized.

The many hours spent in the Despacho 29 helped me build a solid friendship with most of my colleagues that has turn out to be indestructible despite the time and the distance. During my PhD, I have experienced the happiest moments in and outside the office with Laura, Patri and Rafa. As befits people of their age, they have always offered me good advice and lovely support. I would like to thank Laura for showing me the coolest side of Geneva and all the extraordinary days at her apartment where the best breakfast in town is served. With Patri I have drank countless coffees, beers and wines (thanks to maestro Nicolás) all of them seasoned with her extraordinary sense of humor. I owe her a visit to Zürich and Santa María del Águila a.k.a La Aldeílla. The Granada Sound weekend and the Holi Festival with Rafa were two of the funniest parties of my life and the most productive ones in terms of meeting random people. I have scientifically grown up in parallel to Pablo and Juan Carlos that have always been there to solve doubts, share a room in conferences and whose future will be certainly bright.

The very first day of my undergraduate studies in Granada I met Alodía, Emilio, Ernesto, Pablo, Pedro and Rafa. The day after we went together to the Mae West club and just one month later we decided to go on a trip to Bilbao with the incorporation of an unknown girl called Fátima. Since then our lively friendship has travelled all over the globe. A year later and due to a nuclear physics test, Dani and Rafa appeared in my life and started the tradition of "Martes Locos". We have shared wonderful evenings, not only in Granada, but also in different parts of Belgium, Amsterdam, Munich and Barcelona. I could not agree more with the quote "The friends you make in college are friends you'll have for life".

Further, two engineers decisively contributed to enlarge the amount of good memories I have from this time in Granada. Meeting Pedro, a passionate about concrete, beer and hydraulic dams in equal parts, always ensured that my walks back home never lacked a smile on my face. Pablo is one of the sharpest minds I have ever met and is, was and always will be a fundamental person in my life.

Traveling 2000 km north I encountered the SMASHIES crew that welcomed me with the warmest hospitality. I was quickly involved in group dynamics that not only included meetings but also social activities, my favorite being the short walk through the light rain. It was a pleasure to form the spanish gang with Juan, who I expect to meet again in the American adventure, and Laura with whom I expect to share more bus trips to Bielefeld. Thanks to Anna and Jan for providing me good quality coffee together with interesting conversations and helping in the decision-taking process.

Outside the office I have to thank the "Family am Main" members for being always willing to make plans, go out and have some fun no matter the day of the week, nor the weather. Certainly the German experience would have not been the same without my flatmate Basti who first got me out of a real dungeon and later took care of my diet, social life and knowledge of experimental physics.

Going back to my roots and my hometown, there are three guys that have walked by my side the last 20 years of my existence: Anselmo, Jesús and José Antonio. A little bit later joined the group Borja, Chema, Karina, Marina and Mila and next Edu and Zulema completed the "Ajolotes" family. We have experienced almost everything together and keeping them as close friends is one of the things I'm most proud of in this life.

Another common say is that "Friends are the family that one chooses". In my case, I bankrupted the lucky roulette that selects for you a family. Most of the good things I do in this life will be due to the exceptional values that my brother, mother and father instilled in me. Their tireless support and love have easily crossed borders. Thanks for the music festivals, the fun trips together and for making me feel at home just with a text message or a video call. Don’t forget that New York is waiting for the Soto Ontoso family to land.

I would like to end thanking not a person, nor an animal but a city whose streets and clubs have lead me to experience the best moments of my life with most of the people that have appeared in these lines. Thanks for existing, Berlin.

This work was partially supported by a Helmholtz Young Investigator Group VH-NG-822 from the Helmholtz Association and GSI, the Helmholtz International Center for the Facility for Antiproton and Ion Research (HIC for FAIR) within the framework of the Landes-Offensive zur Entwicklung Wissenschaftlich-Oekonomischer Exzellenz (LOEWE) program launched by the State of Hesse, a FP7-PEOPLE-2013-CIG Grant of the European Commission, reference QCDense/631558, and by Ramón y Cajal and MINECO projects reference RYC-2011-09010 and FPA2013-47836.

Contents

Introducción

La física de partículas está viviendo su edad dorada: experimentos como el Relativistic Heavy Ion Collider (RHIC) o el Large Hadron Collider (LHC) están generando petabytes de datos altamente precisos. Pese al enorme progreso realizado en los últimos años, la descripción teórica de la estructura interna del protón continúa siendo uno de los grandes problemas sin resolver en la física de las interacciones elementales. Además de ser una cuestión acuciante y fundamental *per se*, precisar cuáles son los grados de libertad relevantes y sus propiedades tiene profundas implicaciones en diversos aspectos fenomenológicos de las colisiones a altas energías.

La Cromodinámica Cuántica (QCD) se ha consolidado como la teoría cuántica de campos relativista que describe la interacción fuerte. Sus grados de libertad fundamentales son los quarks y los gluones, también llamados globalmente *partones*. Tanto quarks como gluones son portadores de la carga que caracteriza a las partículas que interaccionan fuertemente, esto es, el color. En aquellos procesos físicos en los que el momento transferido es grande, como en Drell-Yan o producción de jets, el éxito de QCD a la hora de describir las observaciones experimentales con un alto grado de precisión es notable. Esto es posible gracias a una propiedad única de la teoría conocida como *libertad asintótica* por la cual la constante de acoplamiento fuerte disminuye a pequeñas distancias o, de manera equivalente, cuando el momento transferido es alto. En estos casos, es posible aplicar métodos perturbativos para calcular las correspondientes secciones eficaces a nivel partónico a cualquier orden.

Sin embargo, el estudio de los constituyentes fundamentales de la materia está dominado por fenómenos de largo alcance, esto es, procesos en los que el momento transferido es pequeño. En este caso la constante de acoplamiento es lo suficientemente grande como para que el cálculo de amplitudes de dispersión quede fuera del dominio de validez de la teoría de perturbaciones. Asimismo, QCD no es capaz de dar una explicación fundamental al hecho de que no se observen quarks y gluones libres en la naturaleza sino combinaciones de ellos en estados neutros de color que llamamos hadrones. Este fenómeno se conoce como *confinamiento* y, a día de hoy, es considerado una conjetura apoyada por observaciones experimentales. Uno de los grandes retos de la física actual es demostrar formalmente esta propiedad de QCD.

Pese a la imposibilidad de describir la estructura hadrónica mediante teoría de perturbaciones en QCD, sí es posible, mediante este marco teórico, determinar su evolución al variar la energía y el momento de la colisión. El paso inicial consiste en medir la estructura del nucleón a una determinada escala de momento y energía. Esta información constituye la condición inicial de las ecuaciones de evolución. Sin duda alguna, los experimentos de dispersión son los microscopios más adecuados para desvelar la estructura básica de las partículas partícipes en la colisión. Siguiendo con la analogía óptica, la resolución de dichos experimentos resulta ser inversamente proporcional a la energía de la colisión, $\Delta r \sim 1/\sqrt{s}$. En concreto, los experimentos de dispersión profundamente inelástica proporcionan acceso a la dependencia en momento longitudinal, x, del hadrón en cuestión, lo que se conoce como función de distribución de partones (PDFs). Los parámetros libres de las PDFs son ajustados a los datos a una determinada escala y su evolución se calcula resumando las correcciones cuánticas que aparecen al variar el momento y la energía de la colisión mediante las ecuaciones del grupo de renormalización. Sin embargo, una caracterización completa del hadrón requiere información tanto en espacio de momentos como de coordenadas. El objeto teórico que contiene ambas dependencias se conoce como distribución de Wigner. Hasta el momento no se ha diseñado ningún experimento capaz de medir la misma. Desde el punto de vista teórico, como se discute en el Capítulo 2 de esta tesis, en su tratamiento aparecen muchas complicaciones como la no-universalidad o la invalidación de los teoremas de factorización. Por lo tanto, en general, la descripción de la estructura hadrónica requiere inevitablemente el uso de herramientas fenomenológicas que combinen modelos teóricos y observaciones experimentales.

La estructura del protón es un elemento relevante en prácticamente cualquier observable en colisionadores hadrónicos tal y como indica la Fig. 1. Actualmente, un tema candente que está siendo investigado intensivamente es la posibilidad de crear gotas de quark gluon plasma (QGP) no solo en colisiones de iones pesados sino también en sistemas no tan densos como en interacciones protón-protón de alta multiplicidad en en LHC. El debate ha sido impulsado por la observación experimental de patrones muy similares en los diferentes sistemas de colisión en el LHC (p+p, p+Pb y Pb+Pb) en los coeficientes armónicos de flujo: una de las pruebas por excelencia de la formación de QGP que es especialmente sensible a la geometría de la colisión inicial.

Otro campo en el que la estructura del protón juega un papel esencial es la descripción de las interacciones múltiples entre partones, un mecanismo fundamental a la hora de entender una colisión hadrónica a energías del LHC, en generadores de eventos Monte Carlo como PYTHIA. Además, una caracterización precisa de las interacciones múltiples entre partones es vital a la hora de controlar el ruido que constituye QCD en búsquedas más allá del Modelo Estándar cuyas señales son débiles.

Por tanto, una caracterización detallada de los grados internos de libertad del protń es un ingrediente crucial del programa de física del LHC. La construcción de futuras instalaciones como el Electron Ion Collider o el LHeC proporcionará una mayor comprensión de la estructura de los hadrones, a través de mapas multidimensionales sobre las distribuciones de partones en el espacio, momento y espín.

Esta tesis se organiza del siguiente modo. En el Capítulo 2 se presentan los elementos básicos del formalismo de QCD y la descripción de la estructura hadrónica en términos de la distribucińn de Wigner y otras funciones que se obtienen a partir de ella integrando (PDFs, distribuciones de momento transverso, distribuciones de partones generalizadas y factores de forma). El objetivo fundamental de esta tesis es caracterizar la estructura transversa del protón. Para ello, hemos analizado los datos de colisiones elásticas protón-protón a energías de ISR y LHC ($\sqrt{s} = 62{,}5$ GeV y 7 TeV, respectivamente). Los resultados de este análisis a $\sqrt{s} = 7$ TeV, presentados en el Capítulo 3, revelan una inesperada propiedad de las interacciones hadrónicas denominada el efecto *hollowness* en la literatura. Proponemos una realización microscópica de esta propiedad desde un punto de vista geométrico mediante el formalismo de Glauber. La descripción del efecto *hollowness* en nuestro modelo es capaz de restringir notablemente la distribución transversa de los grados de libertad subnucleónicos en el protón. Concretamente, hace inevitable la introducción de correlaciones espaciales: algo no considerado en la literatura hasta la fecha. En el Capítulo 4 se expone una amplia discusión acerca de la posible creación de QGP en sistemas de colisión pequeños (p+p, p+Pb). Continuando con estas ideas, el Capítulo 5 está dedicado al estudio de las implicaciones que las correlaciones en el protón tienen en las propiedades del estado inicial en interacciones protón-protón a muy alta energía en el contexto de la física del QGP. Dado el drástico impacto que la inclusión de correlaciones espaciales entre los constituyentes del protón tiene en la descripción del efecto *hollowness*, hemos desarrollado un generador de eventos Monte Carlo que implementa nuestro modelo de la geomtría inicial de la colisión. Todas las propiedades estudiadas son sensibles a la presencia de dichas correlaciones. En especial, su impacto en los cumulantes simétricos ayuda de manera decisiva a reproducir aspectos de las observaciones experimentales. Finalmente, en el Capítulo 6 se exponen las conclusiones de este trabajo.

Ready to start;
Arcade Fire.

Introduction

Particle physics is living its golden age: petabytes of high precision data are being recorded at experimental facilities such as the Relativistic Heavy Ion Collider (RHIC) and the Large Hadron Collider (LHC). Despite the outstanding theoretical progress achieved in the last years the complete understanding of the internal structure of protons remains as one of the most challenging unsolved problems in the physics of elementary interactions. Besides its fundamental interest, pinning down the relevant degrees of freedom (d.o.f) and their properties has profound implications in several phenomenological aspects of high-energy collisions.

Quantum Chromodynamics (QCD) is the, firmly established, relativistic quantum field theory that describes the strong interaction. Its fundamental degrees of freedom are quarks and gluons, collectively called *partons*, both of them carriers of the charge of strongly interacting particles, namely color. In physical processes where the momentum transfer is large, such as Drell-Yan or jet production, the success of QCD in consistently describing the experimental data, up to a high degree of accuracy, is remarkable. This is due to a salient property of the theory, *asymptotic freedom*, by which the strong coupling diminishes at small distances or, equivalently, large momentum transfers, allowing the use of perturbative methods to compute the corresponding partonic cross section to any order. The study of the fundamental constituents of matter is, however, dominated by long-range phenomena, i.e. small momentum transfer, where the strong coupling is large making the calculation of amplitudes for scattering processes beyond the scope of usual perturbation theory. Further, QCD is not able to provide a fundamental explanation to the fact that isolated quarks and gluons are not observed in nature but only combinations of them in colourless bound states called *hadrons*. This is known as *confinement* and it is established as a conjecture supported by experimental facts. To prove it in QCD is still a challenge that hasn't been met.

Thus, hadron structure is not computable within perturbative QCD although its evolution when varying the energy and momentum of the collision can be successfully determined by this framework. The strategy consists in measuring the nucleon structure at a given energy and momentum scale and afterwards use evolution equations. For that purpose, scattering experiments are the most suitable microscopes to unveil the underlying structure of the colliding particles. The resolution of these experiments is inversely proportional to the collision energy, $\Delta r \sim 1/\sqrt{s}$. Deep-inelastic scattering experiments are suitable to measure the longitudinal momentum dependence of the hadron in terms of Parton Distribution Functions (PDFs). Once the PDFs are fitted to data at a given scale their evolution is computed by resuming quantum corrections emerging from the variation of energy and momentum scales of the collisions via renormalization group equations. However, a full characterization of a hadron would require both momentum and spatial information, the so-called Wigner distribution. So far obtaining this information experimentally has not been achieved. From a theoretical point of view several complications arise such as non-universality and breaking of factorization theorems as will be discussed in Chapter 2. Then, in general, the description of the hadron structure relies on phenomenological tools that require theoretical modelling constrained by experimental data.

The proton structure is a relevant input in barely any observable in hadronic colliders as sketched in Fig. 1. Currently one of the subjects undergoing intense study is the possibility that droplets of quark-gluon plasma (QGP) are being created not only in heavy ion collisions but in more dilute systems such as high-multiplicity proton-proton interactions at the LHC. This is a data driven debate as it is rooted in the similar patterns observed across the different collision systems at the LHC (p+p,p+Pb and Pb+Pb) in the flow harmonic analyses: one of the golden probes of QGP formation specially sensitive to the initial collision geometry.

Another domain in which the proton structure plays a central role is the description of multi-parton interactions, the mechanism that dominates the underlying event at LHC energies, in Monte Carlo event generators such as PYTHIA. A precise description of multi-parton interactions is essential to control the QCD background in Beyond the Standard Model searches where the signal is rather weak.

All in all, a detailed characterization of the internal degrees of freedom of the proton is a crucial ingredient of the physics program of the LHC. The construction of future facilities such as the Electron Ion Collider or the LHeC will yield much greater insight into the nucleon structure, by facilitating multi-dimensional maps of the distributions of partons in space, momentum and spin.

This thesis is organized as follows. In Chapter 2, the basics of the QCD formalism and the description of hadron structure in terms of Wigner distributions and its integrated quantities (PDFs, Transverse Momentum Distribution, Generalized Parton Distributions and Form Factors) are presented. Being the main goal of this thesis to characterize the transverse structure of the proton, an analysis of elastic scattering data in p+p interactions at ISR and LHC energies ($\sqrt{s}=62.5$ GeV and 7 TeV, respectively) is performed in Chapter 3. The result of this analysis at $\sqrt{s}=7$ TeV reveal unexpected properties of hadronic interactions dubbed as *hollowness* effect in the literature. We provide a microscopic realization of this feature in a geometric picture such as the Glauber formalism. The description of the hollowness effect within our framework put stringent constraints on the transverse distribution of subnucleonic d.o.f of the proton. More concretely, it requires the introduction of spatial correlations: a unique feature of our model with respect to others in the literature. In Chapter 4, the timely debate on whether QGP effects are being observed in small collision systems (p+p, p+Pb) is exposed. Following these ideas Chapter 5 is devoted to the study of the implications of a correlated description of the proton on the initial state properties of high energy proton-proton interactions in the context of QGP physics. After realizing the drastic impact of the inclusion of spatial correlations between the constituents of the proton in the description of the hollowness effect, we develop a Monte Carlo event generator to implement our model for the initial geometry of the collision. Finally, in Chapter 6 we present the conclusions and the natural extension of the research lines exposed in this thesis.

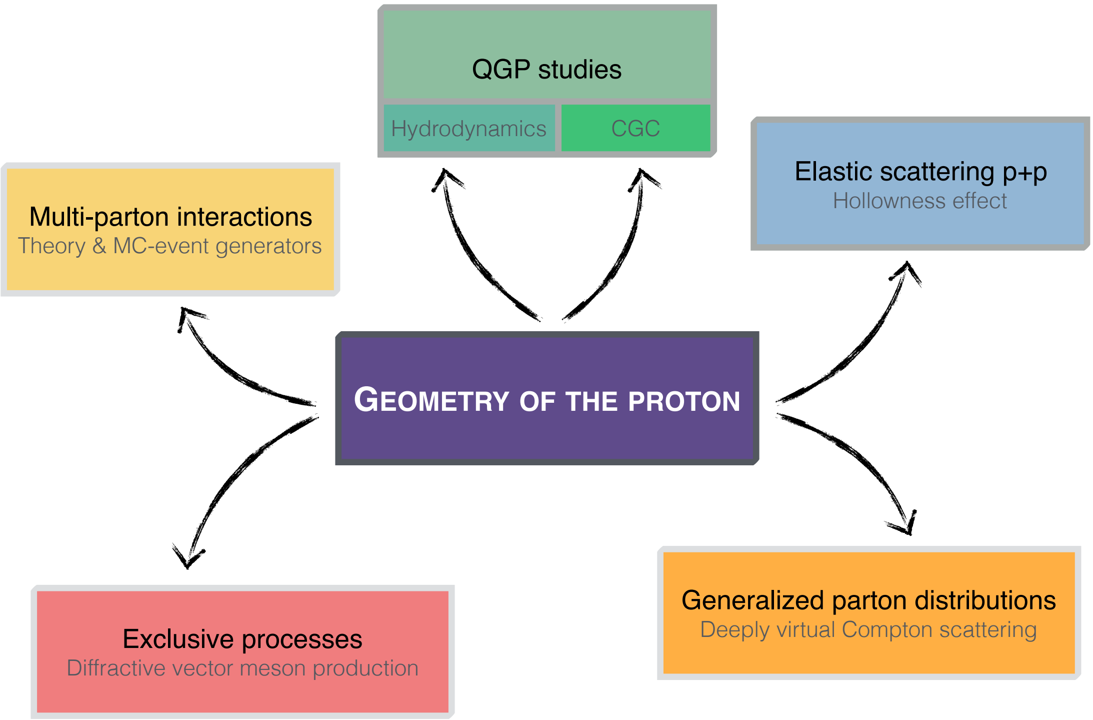

Figure 1: Sketch illustrating some of the different areas of QCD phenomenology where the proton structure is needed as an input. In this thesis we focus on QGP studies and elastic scattering p+p data.

Future starts slow;
The Kills.

1

Quantum Chromodynamics

Currently, we are able to explain physical phenomena based on the existence of four fundamental forces in nature: gravitational, electromagnetic, weak and strong. While gravity is governed by Einstein's general relativity, the other three forces can be described to an excellent accuracy by a quantum field theory of quarks and leptons consistent with Einstein's special theory of relativity and quantum mechanics: the so-called Standard Model (SM). The subject of this work is the study of the internal structure of hadrons and thus we only consider the strong interaction.

This chapter is not intended to be a thorough review of QCD but to emphasize the main features of the theory that will play an important role in the development of subsequent chapters and the more recent experimental results. For a general overview the reader is referred to [1]. Further, a more extended discussion on some relevant experimental facilities namely LHC, RHIC and EIC can be found in Sec. 3.2.

1.1 Basics

Quantum chromodynamics is the theory of the strong interaction. It is a quantum field theory (QFT) with a non-Abelian gauge symmetry group SU($\mathrm{N_c}$), $\mathrm{N_c}=3$, describing the interactions between quarks and gluons. Quarks are point-like, massive, spin-1/2 particles and are represented by a four-component Dirac spinor Ψ. Gluons, in turn, are massless, spin-1 bosons and are represented by four-vector gauge fields A^a_μ. The additional degree of freedom associated with this gauge group is referred to as *color*. The classical Lagrangian density of QCD is:

$$\mathcal{L} = \sum_f \overline{\Psi}_{fi}(\mathrm{i}(D^\mu)_{ij}\gamma_\mu - m\delta_{ij})\Psi_{fj} - \frac{1}{4}F^a_{\mu\nu}F^{a\ \mu\nu} \tag{1.1}$$

with the gluon field strength tensor

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g f^{abc} A^b_\mu A^c_\nu \tag{1.2}$$

and the covariant derivative

$$(D_\mu)_{ij} = \delta_{ij}\partial_\mu - \mathrm{i} g t^a_{ij} A^a_\mu \tag{1.3}$$

where g is the strong coupling constant, t^a are the SU(3) generators in the adjoint representation ($a = 1, ..., 8$) and f^{abc} are the structure constants of the group that determine its Lie Algebra: $[t^a, t^b] = \mathrm{i} f^{abc} t^c$. In Eq. (1.1), the labels (i, j) refer to the colour index of the (anti-)quark that belongs to the fundamental representation of SU(3), $(i, j) = 1, 2, 3$. Additionally, the flavor index, f, of (anti-)quarks can take six possible values $f = u, d, s, c, b, t$ and the QCD Lagrangian is flavor diagonal. Notice that the term $g f^{abc} A^b_\mu A^c_\nu$ in Eq. (1.2), that accounts for gluon self interactions, emerges from the non-Abelian nature of the gauge symmetry and is not present in other field theories such as quantum electrodynamics. This unique feature of QCD allows a quark with a large momentum to radiate gluons that will radiate more gluons themselves and end up in a spray of particles after hadronization, i.e. a jet. In order to extend Eq. (1.1) to the quantum case we would have to consider two additional contributions. First, the gauge-fixing term, that breaks the gauge invariance of the Lagrangian, is needed in any quantum field theory to define the gauge field propagator, in our case the gluon propagator [2]. The inclusion of this term in non-abelian theories goes hand in hand with the Fadaeev-Popov ghosts that allows to preserve gauge symmetry [3]. Details about these terms of the Lagrangian will not be provided here because it is beyond the scope of this thesis.

The nomenclature strong coupling *constant* is slightly misleading as in any renormalizable theory the strength of the interaction actually depends on the energy scale. The running strong coupling constant, to one loop order, is given by [4]

$$\alpha_s(Q^2) = \frac{g^2_{\mathrm{ren}}}{4\pi} = \frac{4\pi}{b \ln(Q^2/\lambda^2)} \tag{1.4}$$

where

$$b = \frac{11\mathrm{N_c} - 2\mathrm{N_f}}{3\mathrm{N_c}} \tag{1.5}$$

and $\mathrm{N_f}$ is the number of active flavors at the energy scale Q^2. λ is a dimensionful energy scale in which the QCD coupling would diverge and it arises due to the renormalization procedure. From Eq. (1.4) the *asymptotic* freedom of QCD can be inferred: when increasing the energy scale of the problem to values much larger than λ, the value of the renormalized coupling constant diminishes and perturbation theory is applicable. In the limit of $Q^2 \to \infty$, $\alpha_s = 0$ and consequently the interaction vanishes. The value of λ gives the approximate scale where non-perturbative effects kick in and therefore perturbation theory breaks down. Thus, $\alpha_s(\lambda)$ does not actually diverge as Eq. (1.4), obtained with perturbative techniques, is no longer valid. Experimentally, see Fig. 1.1, asymptotic freedom is confirmed and λ is found to be ~ 200 MeV, the typical size of a hadron. At this point it is clear that QCD has two different regimes: the one dominated by *soft* phenomena, where perturbative techniques are not applicable, and the *hard* sector where the

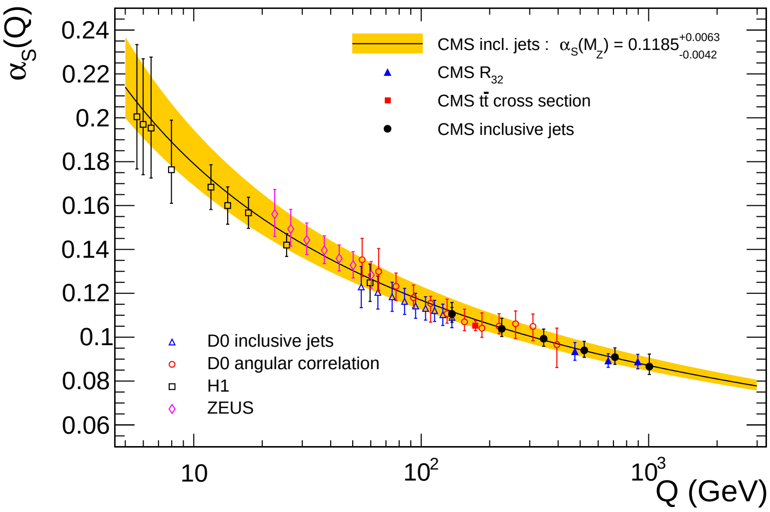

Figure 1.1: The strong coupling α_s and its uncertainty (yellow band) as a function of the momentum transfer Q. Figure extracted from [5].

coupling is small and the full machinery of perturbative QCD can be used. Nevertheless, any process in QCD is sensitive to both of them. Further, although QCD is formulated in terms of quarks and gluons, hadronic states, not elementary quantum fields, are the ones present in real experiments. Thus, a systematic way to relate partonic processes to hadronic states is needed. The way to move forward is to assume that a hadronic cross section can be split into two parts. The total hadronic cross section can be written as

$$d\sigma^{AB} = \sum_{a,b} \int_0^1 \mathrm{d}x_a \mathrm{d}x_b f_{a/A}(x_a) f_{b/B}(x_b) d\hat{\sigma}^{ab} \tag{1.6}$$

That is, a convolution of the parton-level cross section $\mathrm{d}\hat{\sigma}$, that involves only dynamics at short distances i.e. it's computable in perturbative QCD, and the probability of finding such a parton inside the hadron $f(x)$, a non-perturbative object. This procedure relies on a separation of scales. How much information do we need to know about the parton bounded inside the hadron i.e. its longitudinal momentum, transverse position etc., depends on the physical process we are looking at. In the following Sections a rather casual review on the different quantities describing the distribution of partons inside the proton will be presented.

1.2 Wigner distributions

The ultimate understanding of the partonic structure of the nucleon can be gained by means of joint position-and-momentum densities such as the Wigner distributions [6]. To facilitate the discussion we shall start by their definition in quantum mechanics, as it is done in [7]. In a one-dimensional quantum mechanical system with wave function $\psi(x)$, where x represents the spatial coordinate, the Wigner distribution is defined as

$$\mathcal{W}(x,p) = \int \mathrm{d}\eta e^{\mathrm{i}p\eta}\psi(x-\eta/2)^*\psi(x+\eta/2). \tag{1.7}$$

Due to Heisenberg's uncertainty principle Eq. (1.7) doesn't have a probabilistic interpretation and for arbitrary values of x and p it is not positive definite. However, when integrating out the coordinate x one recovers the genuine momentum space density, $|\psi(p)|^2$, and, equivalently, the coordinate space density, $|\psi(x)|^2$ is obtained by p integration. Then, the Wigner distribution contains all the information available in the wave function. Our aim is to extend this picture to the description of quarks in a quantum field theory. First, one has to promote wave functions to operators, in this case, spinors:

$$\mathcal{W}(\vec{r},k) = \int \mathrm{d}^4\eta e^{\mathrm{i}k\cdot\eta}\overline{\psi}(\vec{r}-\eta/2)\psi(\vec{r}+\eta/2) \tag{1.8}$$

where $\vec{r}$ is the position of the quark with respect to the centre of the proton. k is its four momentum written in light-cone coordinates as $k = (k^+, k^-, \mathbf{k}_\perp)$ conjugate to the space-time separation z[1]. The definition of light-cone coordinates is given in Appendix B. Further, as quarks are spin-$1/2$ particles it is needed to include a Dirac matrix, Γ,

$$\mathcal{W}(\vec{r},k) = \int \mathrm{d}^4\eta e^{\mathrm{i}k\cdot\eta}\overline{\psi}(\vec{r}-\eta/2)\Gamma\psi(\vec{r}+\eta/2) \tag{1.9}$$

that selects the polarization of the quark inside the proton. For instance, in the case of an unpolarized quark one has $\Gamma = \gamma^+ = (\gamma^0+\gamma^z)/\sqrt{2}$ [8]. From Eq. (1.9) it can be easily deduced that gauge invariance is broken as we are evaluating the two fermionic fields at different space-time points. More concretely, any bi-local operator product such as $\overline{\psi}(y)\psi(x)$ is not invariant under gauge transformations $\mathcal{U}$

$$\begin{aligned}
&\overline{\psi}(x) \to \mathcal{U}^\dagger(x)\psi(x)\\
&\psi(x) \to \mathcal{U}(x)\psi(x)\\
&\qquad \Rightarrow \overline{\psi}(y)\psi(x) \to \psi(y)\mathcal{U}^\dagger(y)\mathcal{U}(x)\psi(x) \neq \overline{\psi}(y)\psi(x).
\end{aligned} \tag{1.10}$$

This difficulty should be overcome as the Wigner distribution could be extracted from experimental measurements and, consequently, it must be gauge invariant. To fix this problem, we have to insert an operator, $\mathcal{O}_{[x,y]}$, in between the quark fields with the following gauge transformation properties

$$\mathcal{O}_{[y,x]}\psi(x) \to \mathcal{U}(y)\left(\mathcal{O}_{[y,x]}\psi(x)\right) \tag{1.11}$$

that would lead to the bi-local operator product supplemented with $\mathcal{O}_{[x,y]}$ being gauge invariant

$$\begin{aligned}
\overline{\psi}(y)\mathcal{O}_{[y,x]}\psi(x) &\to \overline{\psi}(y)\mathcal{U}^\dagger(y)U(y)\left(\mathcal{O}_{[y,x]}\psi(x)\right)\\
&\Rightarrow \overline{\psi}(y)\mathcal{O}_{[y,x]}\psi(x).
\end{aligned} \tag{1.12}$$

---

[1] Bold face is used to denote two dimensional vectors.

This condition is fulfilled by the so-called gauge link, $\mathcal{G}$, [9]

$$\mathcal{G}_{[x,y]} = \mathcal{P}\exp\left[-ig\int_x^y d\xi^\mu A_\mu^a(\xi)T_a\right] \tag{1.13}$$

where $\mathcal{P}$ reflects the need to order the exponential in the gauge group generators T_a given the non-Abelian nature of QCD i.e. they do not conmute. The definition given by Eq. (1.13) is identical to the one of a Wilson line, $\mathcal{U}_{[x,y]}$, that describes the propagation of a high energy parton through a background color field. The connection between these two concepts and a more intuitive interpretation of Eq. (1.13) in terms of multiple gluon exchanges between the struck quark and the fragments of the proton after the scattering will be given in Sec. 1.2.1.

Thus, including all the elements discussed in Eqs. (1.9-1.13) we obtain the complete expression for the Wigner operator

$$\mathcal{W}^{\mathbf{q}}(\vec{r},k^+,\mathbf{k}_\perp) = \int \mathrm{d}^3\eta e^{-\mathrm{i}k^+\eta^-+\mathrm{i}\mathbf{k}_\perp\eta_\perp}\overline{\psi}(\vec{r}-\eta/2)\Gamma\mathcal{G}_{[\vec{r}-\eta/2,\vec{r}+\eta/2]}\psi(\vec{r}+\eta/2) \tag{1.14}$$

where we have imposed that both quarks have equal light-cone time η^+ [10] or, equivalently, we have integrated out x^-. This step reflects our ignorance on how to measure the light-cone energy of the quarks involved in high-energy experiments, the perfect playground to resolve the partonic content of the proton and thus be able to measure the Wigner distributions. Note that although the gauge link only depends on the endpoints, fixed by the Wigner operator, there is still the freedom of the choice of the path ξ i.e. how to reach the endpoints. The importance of this fact will be explicitly shown in Sec. 1.2.2. Therefore, the gauge invariant Wigner operator is, in general, path dependent. The last step to construct joint position-momentum distribution from the Wigner operator given by Eq. (1.14) is to compute its expectation value between localized proton states

$$W^q(\vec{r},k^+,\mathbf{k}_\perp) = \langle\vec{R}=0|\mathcal{W}^{\mathbf{q}}(\vec{r},k^+,\mathbf{k}_\perp)|\vec{R}=0\rangle \tag{1.15}$$

or, equivalently, in momentum space

$$\widetilde{W}^q(\xi,\mathbf{\Delta}_\perp,k^+,\mathbf{k}_\perp) \equiv 2m_p\int\frac{\mathrm{d}^3r}{(2\pi)^3}e^{\mathrm{i}\vec{\Delta}\cdot\vec{r}}W^q(\vec{r},k^+,\mathbf{k}_\perp) \tag{1.16}$$

where m_p is the mass of the proton and the spin dependence has been omitted. The definition of a localized proton state is cumbersome given its relativistic nature. For instance, the recoil effect cannot be neglected. The choice of $\vec{R}=0$ is justified by looking at its conjugate variable in the Breit frame (see Appendix A) where the incoming proton has a momentum $P-\Delta/2$ and the outgoing one $P+\Delta/2$ as depicted in Fig. 1.2. The fact that the momentum transferred to the proton is reversed after the scattering is equivalent to set $\vec{R}=0$ in coordinate space. For consistency, the rest of the calculations are done in this reference frame. Although we are referring to the quark Wigner distribution the extension to the antiquark and gluonic cases is straightforward. Explicit expressions can be found in [11].

To sum up, the Wigner distribution for a quark, given by Eq. 1.15 is a 6 dimensional function: 3 spatial and 3 momentum coordinates. So far the challenge of designing

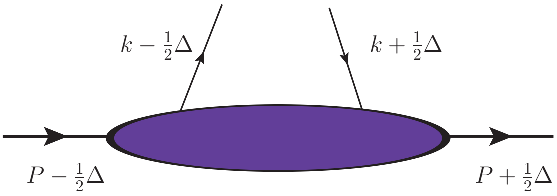

Figure 1.2: Kinematics for the quark Wigner distribution in terms of k, the quark momentum, P, the proton momentum, and Δ, the momentum transferred to the proton.

experimental observables to measure it has not been met. However, by integrating with respect to different variables, see Fig. 1.3, we can generate distributions that contain less information about the proton but that, in some cases, can be measured. The purpose of next Sections is to exploit the concepts that have been presented above to examine the most relevant distribution functions for high-energy collisions, namely PDFs, Transverse Momentum Distributions (TMDs) and Generalized Parton Distributions (GPDs), both from an experimental and a theoretical point of view.

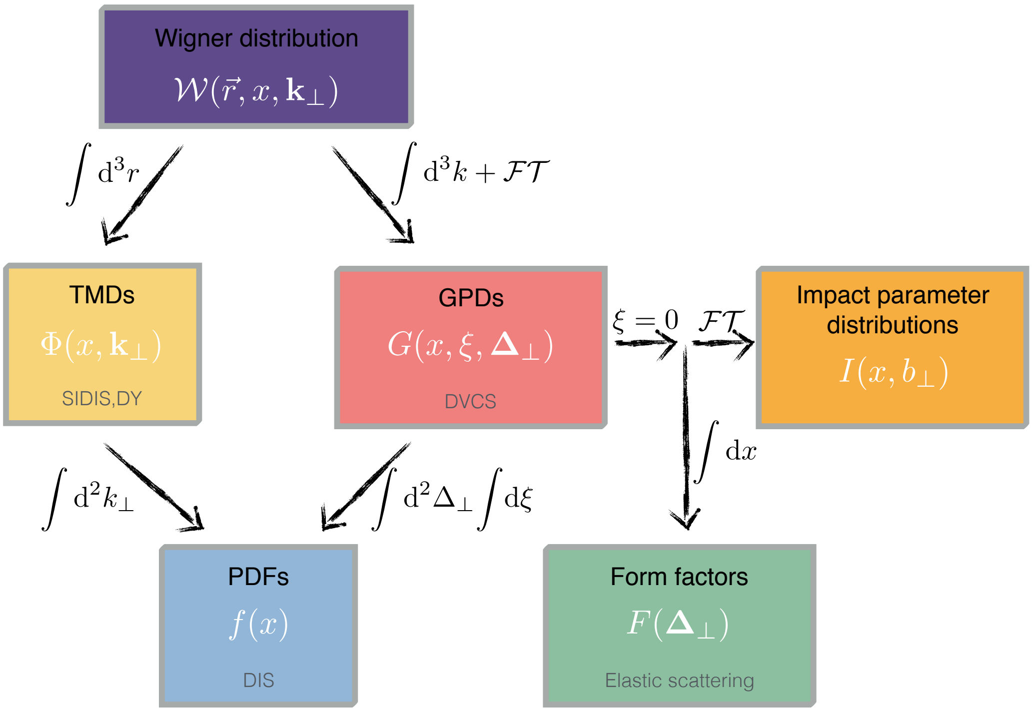

Figure 1.3: Sketch illustrating some of the quantities that can be derived from the Wigner distribution and the experimental processes where they can be measured. $\mathcal{FT}$ denotes a Fourier transform from momentum to coordinate space.

1.2.1 PDFs and DIS

The most familiar element of Fig. 1.3 are the parton distribution functions that contain information about the longitudinal momentum structure of the proton. Their formal definition, for the unpolarized quark case, in the same language as in Sec. 1.2 is

$$f^q(x) = \int \frac{\mathrm{d}\eta^-}{2\pi} e^{-ixP^+\eta^-} \langle P|\overline{\psi}(0,\eta^-,\mathbf{0}_\perp)\gamma^+\mathcal{G}_{[\eta^-,0]}\psi(0,0,\mathbf{0}_\perp)|P\rangle\Big|_{\eta^+=\eta_\perp=0} \tag{1.17}$$

where P^+ is the momentum of the proton and $k^+ = xP^+$[2]. Thus, PDFs describe the probability of finding a parton with longitudinal momentum fraction x of its parent hadron. In this case the gauge link is greatly simplified

$$\mathcal{G}^{\mathrm{DIS}}_{[(0,\eta^-,\mathbf{0}_\perp),(0,0,\mathbf{0}_\perp)]} = \mathcal{P}\exp\left[-ig\int_0^{\eta^-} \mathrm{d}\xi^- A_a^+(0,\xi^-,\mathbf{0}_\perp)t^a\right] \tag{1.18}$$

and, in particular, it can be set to one by working in the light-cone gauge $A^+ = 0$. This can also be written as the product of two longitudinal Wilson lines:

$$\mathcal{G}_{[(0,\eta^-,\mathbf{0}_\perp),(0,0,\mathbf{0}_\perp)]} = \mathcal{U}^L_{[\eta^-,\infty;\mathbf{0}_\perp]}\mathcal{U}^L_{[\infty,0;\mathbf{0}_\perp]} \tag{1.19}$$

with

$$\mathcal{U}^L_{[a,b;x_\perp]} = \mathcal{P}\exp\left[\mathrm{i}g\int_a^b dx^- A_a^+(0,x^-,x_\perp)t^a\right] \tag{1.20}$$

as represented in Fig. 1.4. An important comment is in order at this point: up to now we have not assumed anything about the physical process, in which this quark is involved. The fact that the formal operator definition of the PDF does not depend on the gauge link is directly related with the concept of *universality* i.e. the parton distributions are properties of the hadron itself, and not of the collision. This endows QCD with a high predictive power: the PDF is an intrinsic property of the quark and if we measure it in some process, e.g. electron DIS, we can reuse it in another experiment such as proton-proton interactions. Before exploring the phenomenological issues of PDFs there is still one relevant aspect to be clarified about the gauge link. We introduced $\mathcal{G}_{[x,y]}$ somehow ad-hoc in order to make the Wigner operator gauge-invariant. However its physical meaning remains blurry. The fact that $\mathcal{G}_{[x,y]}$ shares the same expression as the one for a Wilson line is not an artificial coincidence. The Wilson line attached to $\psi(x)$ in Eq. (1.17) resums all soft gluon exchanges between the struck quark and the color gauge field of the proton. Qualitatively, the parton that suffers the hard process can rescatter off the proton. These interactions can happen before or after the hard scattering depending on the physical process. The path-ordering ensures that the integrations are performed in the same order in which the interaction occur. The type of diagrams that contribute to the gauge link are shown in Fig. 1.5. An explicit calculation of the left process is presented in Appendix C. To finish with this discussion it should be emphasized that when two quarks

---

[2] In the polarized case, two more PDFs appear, namely the helicity distribution function, denoted as $g_L(x)$, obtained by setting $\Gamma = \gamma^+\gamma^5$ and the transversity distribution $h(x)$ where $\Gamma = -[\gamma^i,\gamma^+]\gamma^5$.

from different hadrons come into play, factorizing multiple gluon exchanges is not trivial i.e. there could be interference such as one of the quarks rescattering both off its parent hadron and the projectile. It is well beyond the scope of this project to discuss how to overcome these difficulties and also the divergences that arises from the gauge link. An updated discussion on this topic can be found in [12].

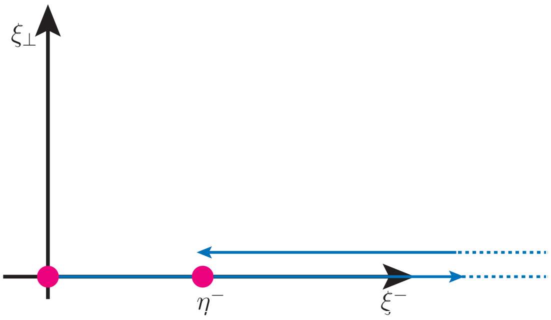

Figure 1.4: Path of the gauge link given by Eq. (1.19). The dotted lines indicate the extension up to ∞.

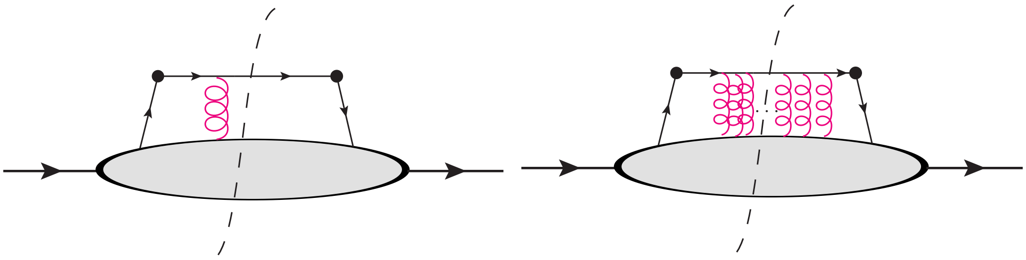

Figure 1.5: Left: One-gluon exchange between the parton and the background color field of its parent hadron. Right: n-gluon exchange as accounted by the Wilson line. The black dots represent any hard interaction that the struck quark may suffer. The dotted line represents the cut that distinguishes the scattering amplitude and its conjugate.

It has been stressed that although parton distribution functions are not computable from first principles in QCD as they are dominated by the non-perturbative regime, they are well defined objects, see Eq. (1.17). Thus, the only way out to access parton distribution is via experimental measurements. Inclusive processes where a hard probe struck the parton inside the proton, breaks it and the fragments are not measured are the adequate playground. Deep ($Q^2 \gg m_p^2$) inelastic ($m_X^2 \gg m_p^2$) electron-proton scattering, shown on the left side of Fig. 1.6, is the cleanest way to test the hadron structure as the electromagnetic interaction of leptons is well understood. In this reaction, only the electron is detected and the final states are integrated out. Consequently the sensitivity to the transverse momentum of the struck quark is lost and this process is ideal to explore the longitudinal momentum dependence of partons inside hadrons. The assumption that the photon interacts with only one constituent of the proton can be justified by looking at the timescales. For that purpose it is convenient to use a reference frame in which the hadron moves at high velocity: due to Lorentz time-dilation the time scale characteristic of interactions between the partons is far larger than that characteristic of the interaction with the virtual boson ($\Delta T = 1/Q$). Therefore, the incoming photon scatters incoherently and instantaneously off a single parton. The differential cross section for the process on the right side of Fig. 1.6 in the fixed-target frame is given by [13]:

$$E'_k \frac{\mathrm{d}^3\sigma}{\mathrm{d}k'^3} = \frac{2}{s - m_p^2} \frac{\alpha^2}{Q^4} L^e_{\mu\nu} W^{\mu\nu} \tag{1.21}$$

where α is the fine structure constant. $L^e_{\mu\nu}$ is the leptonic tensor given by a straightforward calculation in QED; its expression is [13]

$$L^e_{\mu\nu} = 2(k'_\mu k_\nu + k'_\nu k_\mu - g_{\mu\nu}(k' \cdot k - m^2)). \tag{1.22}$$

The hadronic tensor $W^{\mu\nu}$ contains both the virtual photon-quark scattering and the non-perturbative information about the proton, see Fig. 1.7. It can be expressed, as expected, in terms of the parton distribution functions as defined by Eq. (1.17) [14]

$$W^{\mu\nu} \sim \sum_q e_q^2 \mathrm{Tr}(f^q(x) \gamma^\mu \gamma^+ \gamma^\nu). \tag{1.23}$$

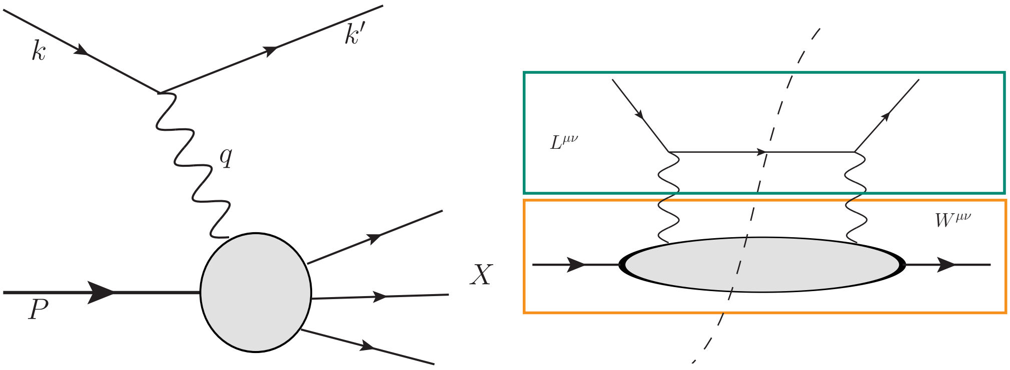

Figure 1.6: Left: Deep inelastic scattering: $e(k)p(P) \to e(k')X$. Right: Sketch representing the manifest separation between the leptonic tensor (green box) and the hadronic one (orange box).

An important assumption was made to write the hadronic tensor as in Eq. (1.23): the long-distance dynamics encoded in the parton distributions decouples from the hard-scattering between the photon and the quark. The latter is computed, in this case, at leading order in perturbation theory assuming on-shell massless partons independent of the non-perturbative part of the process. In other words, the differential DIS cross section can be written schematically as

$$\frac{d\sigma^{eP}}{dx} \sim f^q(x) \otimes \frac{d\sigma^{eq}}{dx} \tag{1.24}$$

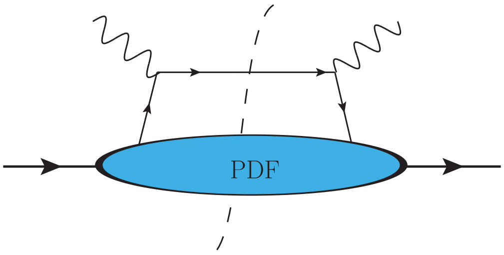

Figure 1.7: The hadronic tensor $W^{\mu\nu}$ decomposed as the γq-scattering and the parton distribution function.

that is, the electron-proton scattering is split into two parts: the photon-quark scattering characterized by a hard scale ($Q^2 \gg \lambda$) and the parton distribution function. This is the main result of the so-called collinear factorization scheme. The meaning of *collinear* becomes clear in the infinite momentum frame in which the proton has a very large momentum in the longitudinal direction: due to Lorentz-contraction the partons are almost parallel to the direction of motion of the proton. From Eq. (1.24) it is unclear how to make the distinction between *hard* and *soft* modes. In other words, the factorization is done at an arbitrary scale μ_F^2 in such a way that the hard part contains $k_\perp^2$ larger than this scale and the non-perturbative part has $k_\perp^2 < \mu_F^2$. Thus,

$$f^q(x) \to f^q(x, \mu_F^2). \tag{1.25}$$

μ_F^2 is called *factorization scale* and its value is usually set to the virtuality of the process Q^2. Certainly, a physical observable cannot depend on the choice of μ_F^2. Fortunately, the μ_F^2-dependence of the parton distribution functions can be calculated perturbatively, provided that μ_F^2 is sufficiently large, in terms of evolution equations. An intuitive insight on the fact that the PDF depends on μ_F^2 or, equivalently, Q^2, can be achieved by considering that the quark struck by the virtual photon may have a history prior to the interaction i.e. it may come from a radiation process or have radiated itself new gluons as shown on the right side of Fig. 1.8. From this point of view, the distributions $f^q(x, Q^2)$ can be interpreted as the probability of finding a quark, where emissions with transverse momentum lower than Q^2 are not resolved [15]. Therefore, to compute the evolution we have to resum all the gluon emissions of the quark up to a scale Q^2.

Consider the emission of a gluon which carries a small fraction $x \ll 1$ of the longitudinal momentum of its parent quark represented in Fig. 1.8. The differential probability for this emission is given by

$$\mathrm{d}P \simeq \frac{\alpha_s C_F}{2\pi} \frac{\mathrm{d}^2 k_\perp}{k_\perp^2} \frac{\mathrm{d}x}{x} \tag{1.26}$$

where $\mathbf{k}_\perp$ and x are the transverse momentum and the fraction of longitudinal momentum of the newly created parton, respectively. There are two types of singularities in this expression: when $x \to 0$ (soft divergence) and when $k_\perp^2 \to 0$ (collinear divergence). However, the soft singularities cancel when the virtual corrections are taken into account [12]. Concerning the collinear divergence, we can understand what is going on, if we realize that

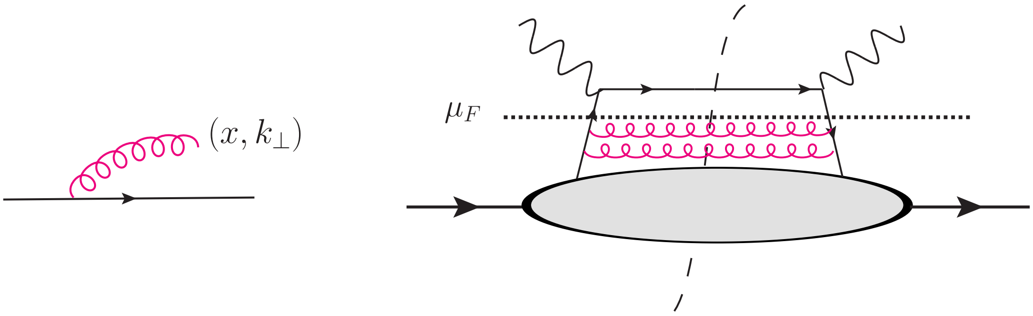

Figure 1.8: Left: a quark radiating a soft gluon with momentum components $(x, \mathbf{k}_\perp)$. Right: the transverse momenta of the gluons are smaller than the factorization scale so they are absorbed in the PDF via DGLAP equations. If $k_\perp^2 > \mu_F^2$ the gluon is a correction to the hard part.

the limit $k_\perp^2 \to 0$ corresponds to a long-range part of the strong interaction, which is not calculable in perturbation theory. The collinear singularities can then be absorbed into a redefinition of the parton densities. Once Eq. (1.26) is integrated over the available phase space region, parametrized by the energy scales of the DIS process (x, Q^2), the probability of gluon emission is going to be enhanced by large logarithms of Q^2 or $1/x$. In order to get reliable predictions, these logarithms have to be resummed to all orders.

The Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) evolution is the most familiar resummation strategy [16, 17, 18]. It provides the evolution of hadron structure to higher values of Q. Such processes are dominated by diagrams where the transverse momenta of the successive emitted partons are strongly ordered,

$$\frac{1}{R_h} \ll k_{\perp 1} \ll k_{\perp 2} \ldots \ll k_{\perp n} \sim Q \tag{1.27}$$

where R_h is the length scale of the order of the hadron size and the longitudinal momenta are less constrained. The $\mathbf{k}_\perp$-ordering gives rise to large logarithms $\ln Q^2$ as the probability of emission of n gluons, using Eq. 1.26, is

$$\mathrm{d}P_n \sim \int_{Q_0}^{k_{\perp 2}} \frac{\mathrm{d}^2 k_{\perp 1}}{k_{\perp 1}^2} \int_{k_{\perp 1}}^{k_{\perp 3}} \frac{\mathrm{d}^2 k_{\perp 2}}{k_{\perp 2}^2} \ldots \int_{k_{\perp n-1}}^{Q} \frac{\mathrm{d}^2 k_{\perp n}}{k_{\perp n}^2} \sim \left[\alpha_s(Q^2) \ln\left(\frac{Q^2}{Q_0^2}\right)\right]^n \tag{1.28}$$

where Q_0 is some initial perturbative scale. In summary, DGLAP resums terms $\alpha_s \ln(Q^2/Q_0^2) \sim 1$ to all orders. The complete DGLAP evolution equations can be written in a compact matrix way that explicitly shows how different components are mixed through evolution:

$$\frac{\partial}{\partial \ln Q^2} \begin{pmatrix} \Sigma(x, Q^2) \\ f^g(x, Q^2) \end{pmatrix} = \frac{\alpha_s(Q^2)}{2\pi} \int_x^1 \frac{\mathrm{d}z}{z} \begin{pmatrix} P_{qq}(z) & 2n_f P_{qg}(z) \\ P_{gq}(z) & P_{gg}(z) \end{pmatrix} \times \begin{pmatrix} \Sigma(x/z, Q^2) \\ f^g(x/z, Q^2) \end{pmatrix} \tag{1.29}$$

where $\Sigma(x, Q^2) = f^q(x, Q^2) + f^{\bar{q}}(x, Q^2)$ and n_f is the number of active flavours. These equations are the analogue to the β function describing the variation of $\alpha_s(Q^2)$ as given by Eq. (1.4). The DGLAP equations have been successfully and intensively tested against experimental data and, together with asymptotic freedom and factorization theorems, provide a fundamental tool for establishing controlled theoretical predictions.

To obtain the parton distributions from data we assume a given structure in the x variable, such as a polynomial function with a given number of free parameters, at a low Q_0^2. The initial parameters are fitted to data once and for all. Then, by using DGLAP equations and the universal nature of the PDFs, our knowledge of the proton at a certain value Q_0^2 can be used for other observables at any higher Q^2. An important clarification should be made. This procedure works as long as no new dynamical effects become relevant in the region where the set of PDFs is applied. For instance, the kinematic region relevant for ultra-high energy neutrino-nucleon interactions is characterized by very small values of x ($x \lesssim 10^{-7}$ for $E_\nu \gtrsim 10^{11}$ GeV) and virtualities of the order of the electroweak boson mass squared, $Q^2 \sim M_{Z,W}^2 \sim 10^4$ GeV^2. Clearly, DGLAP equations are not sufficient to capture all the dynamics as not only $\ln Q^2$ are relevant in this scenario but $\ln 1/x$ too [19].

The HERA experiment at the Deutsches Elektronen Synchrotron (DESY) laboratory in Hamburg can be considered the particle accelerator *par excellence* in the field of DIS. Several fixed target experiments explored the interaction of electron beams with hadronic matter. However, HERA has been the only electron/positron-proton collider ever built and took data from 1997 to 2004. The electrons were accelerated to an energy of 27.5 GeV and the protons up to 920 GeV. All in all, the center of mass energy of these collisions was $\sqrt{s} = 2\sqrt{E_e E_p} = 318$ GeV. The kinematic region in the (x, Q^2)-plane explored by the two main detectors at HERA, H1 and ZEUS, is shown in the left part of Fig. 1.9. Fixed target experiments and recent data taken from interactions distinct to DIS at the LHC are also included. This is the data set used by the NNPDF Collaboration together with

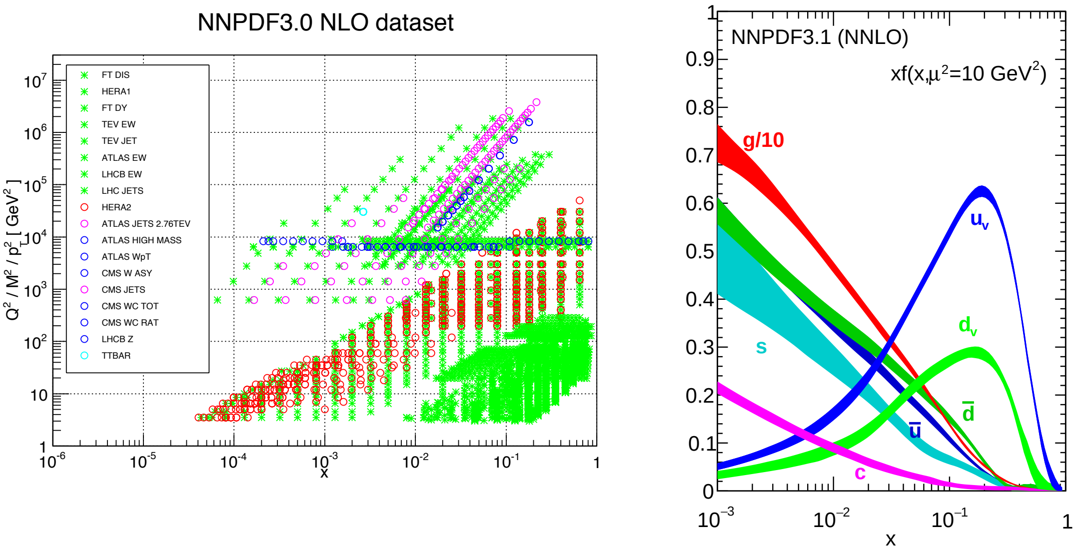

Figure 1.9: Left: Kinematic coverage in the (x,Q^2)-plane of different experiments. Right: The NNPDF3.1 NNLO PDFs, evaluated at $\mu_F^2 = 10$ GeV^2. Figures extracted from [20].

state-of-the-art artificial intelligence techniques to constrain the free parameters of their parametrizations. As an example, in Fig. 1.9 the PDFs of a proton multiplied by x at $Q^2 = 10$ GeV^2 are shown as a function of x. At large x, the two valence quarks, u_v, carry approximately $2/3$ of the total momentum and the d_v quark carries approximately $1/3$ of the total momentum. However, with decreasing x the population of the sea quarks, s, rises. The gluon contribution becomes clearly dominant at small-x. The relation between a large gluonic density at small-x and the applicability of weak coupling techniques will be discussed in Sec. 1.3.

1.2.2 TMDs and SIDIS

The enormous success of the (universality+collinear factorization+DGLAP equations)-paradigm when describing data up to a high degree of accuracy is, probably, one of the major achievements of QCD. However, it turns out to be insufficient when measuring more exclusive processes i.e. where not all the final state is integrated out. In these cases, going beyond the one-dimensional characterization of the proton is mandatory. More concretely, consider a DIS process in which not only the lepton but also one hadron is detected after the scattering. This is known as Semi-Inclusive DIS (SIDIS) and is represented in Fig. 1.10.

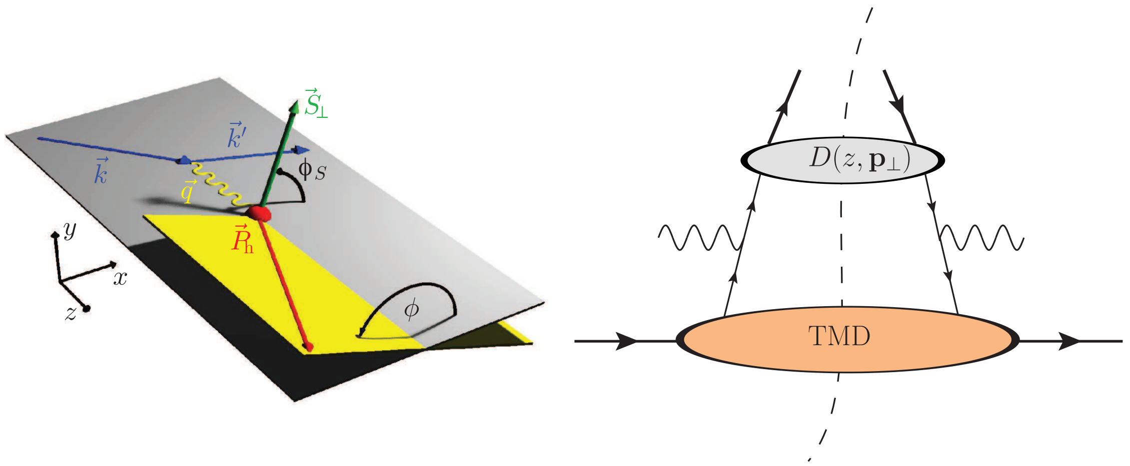

Figure 1.10: Left: Parton level SIDIS kinematics. Figure extracted from [21]. Right: SIDIS factorizes as the convolution of the TMD, the hard interaction with the virtual photon and the fragmentation function.

The measured hadron is directly related with the scattered quark as it originated from its fragmentation. Thus, by measuring the scattered lepton and the momentum of this hadron one can access both the longitudinal and transverse momentum information of the struck quark. Furthermore, the fragmentation function $D(z, \mathbf{p}_\perp)$, neglected in the DIS scenario where the final state was integrated out, depends as well on the transverse momentum of the quark. Following Fig. 1.3, the joint distribution of partons in their longitudinal momentum fraction x and their momentum transverse to the proton direction $\mathbf{k}_\perp$ is given by the Transverse-Momentum Dependent parton densities (TMDs) defined as

$$\Phi(x,\mathbf{k}_\perp) = \int \frac{\mathrm{d}\eta^- \mathrm{d}^2\eta_\perp}{(2\pi^3)} e^{-\mathrm{i}xP^+\eta^- + \mathrm{i}k_\perp\cdot\eta_\perp}\langle P|\overline{\Psi}(0,\eta^-,\eta_\perp)\gamma^+\mathcal{G}_{[(\eta^-,\eta_\perp);(0,\mathbf{0}_\perp)]}\Psi(0,0,\mathbf{0}_\perp)|P\rangle \tag{1.30}$$

in the case where the quark is unpolarized. When considering both the polarization of the quark and of its parent hadron eight independent TMD quark distributions arise. They are summarize in Fig. 1.11. Three of them can be considered as the natural extension of the PDFs ($f^q(x)$,$g_L(x)$ and $h(x)$)[3], while the origin of the other five lies on the vectorial nature of $\mathbf{k}_\perp$ and vanish when a $\mathbf{k}_\perp$-integration is performed.

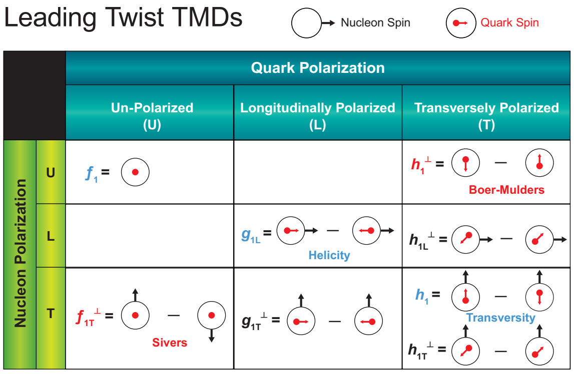

Figure 1.11: Different TMDs classified according to the polarization of the quark and the nucleon. Figure extracted from [21].

A novel structure of the gauge link appears in Eq. (1.30) with respect to the PDF case, Eq. (1.17). The space-time separation between the fields does not lie on the light-cone but has a transverse component. Then, apart from a Wilson line in the longitudinal direction as defined in Eq. (1.20) one in the transverse direction is needed,

$$\mathcal{U}^T_{[x^-;a_\perp,b_\perp]} = \mathcal{P}\exp\left[\mathrm{i}g\int_{a_\perp}^{b_\perp}\mathrm{d}x_\perp A^a_\perp(0,x^-,x_\perp)T_a\right]. \tag{1.31}$$

Following our previous interpretation, inside the TMD we have a resummation of both soft collinear gluons coming from $\mathcal{U}_\mathrm{L}$ and of the transverse ones as this is the role of $\mathcal{U}_\mathrm{T}$. Clearly, the whole gauge link cannot be set directly to one by choosing the light-cone gauge. This will only drop the longitudinal part. At this level this indicates that the TMDs are going to be path-dependent. The complete expression for the gauge link in the case of SIDIS is given by

$$\mathcal{G}^{\mathrm{SIDIS}}_{[(\eta^-,\eta_\perp);(0,\mathbf{0}_\perp)]} = \mathcal{U}^L_{[\eta^-,\infty;\eta_\perp]}\mathcal{U}^T_{[\infty;\eta_\perp,\infty]}\mathcal{U}^T_{[\infty;\infty,\mathbf{0}_\perp]}\mathcal{U}^L_{[\infty,0;0_\perp]} \tag{1.32}$$

and it is represented on the left side of Fig. 1.12. Up to now we have only considered situations (DIS and SIDIS) where the gluon exchanges between the struck quark and the

---

[3] This identification only holds when QCD corrections are neglected.

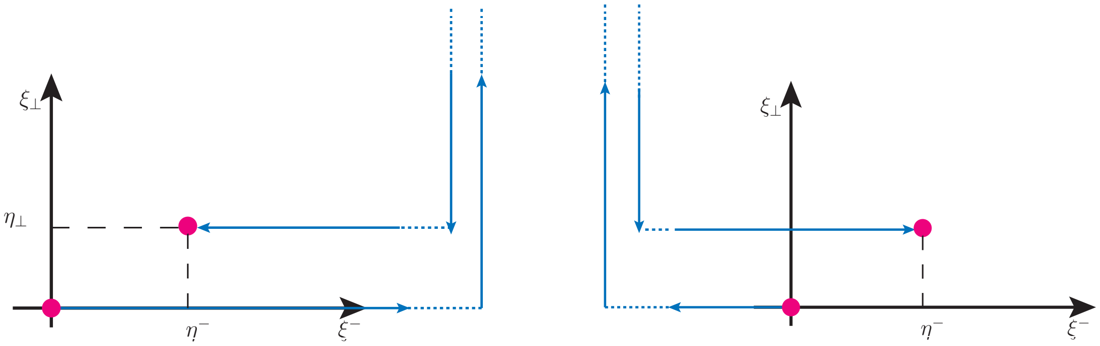

Figure 1.12: Left: Path of the gauge link in SIDIS as given by Eq. (1.32). Right: Path of the gauge link in DY as given by Eq. (1.33).

spectators in its parent hadron occur after the hard scattering i.e. they are final state interactions and the Wilson lines are future pointing (they extend up to $+\infty$). However, this is not always the case. Consider a Drell-Yan process (DY) where quark and antiquark annihilate to form a virtual photon which then decays into a pair of oppositely-charged leptons. Then, the multiple gluon exchanges between the struck quarks and the hadrons take place before the hard scattering and the gauge-link for each participant looks like

$$\mathcal{G}^{\rm DY}_{[(\eta^-,\eta_\perp);(0,\mathbf{0}_\perp)]} = \mathcal{U}^L_{[\eta^-,-\infty;\eta_\perp]}\mathcal{U}^T_{[-\infty;\eta_\perp,\infty]}\mathcal{U}^T_{[-\infty;\infty,\mathbf{0}_\perp]}\mathcal{U}^L_{[-\infty,0;0_\perp]} \tag{1.33}$$

where the Wilson lines are "past pointing" or, in other words, define an initial state interaction as depicted in the right part of Fig. 1.12. By comparing Eqs. (1.32) and (1.33) an important consequence is extracted: TMDs are process-dependent, they are not universal. In particular, this is an important result for the two naively[4] time-reversal odd TMDs from Fig. 1.11 i.e. the ones that do not fulfill

$$\Phi(\vec{k},\vec{P},\vec{S})^* = \gamma^5(\mathrm{i}\gamma^2\gamma^0)\Phi(-\vec{k},-\vec{P},-\vec{S})(i\gamma^{2\dagger}\gamma^{0\dagger}) \tag{1.34}$$

the so-called Sivers ($f_{1T}^\perp(x,\mathbf{k}_\perp)$) [22] and Boer-Mulders functions ($h_1^\perp(x,\mathbf{k}_\perp)$) [23]. According to our discussion the following relation should be realized

$$f_{1T}^\perp(x,k_\perp)\Big|_{\rm SIDIS} = -f_{1T}^\perp(x,k_\perp)\Big|_{\rm DY} \tag{1.35}$$

and the same for $h_1^\perp(x,\mathbf{k}_\perp)$. Together with their non-zero value, the relation given by Eq. (1.35) constitutes a very robust prediction of the TMD formalism stemming from gauge invariance: ($f_{1T}^\perp(x,\mathbf{k}_\perp)$, $h_1^\perp(x,\mathbf{k}_\perp)$) are not universal but exhibit time-reversal modified universality. In order to connect the operator definition of a TMD and the experimental measurement of Eq. (1.35) via DY and SIDIS processes an essential ingredient is required: a factorization theorem in the spirit of Eq. (1.24). In the case of SIDIS the

---

[4] The term *naive* is used to emphasized the fact that the origin of the time-reversal breaking are the gluon exchanges. Without them these distributions would be zero.

type of factorization for the total differential cross section that has been probed to be valid when $\mathbf{p}_\perp \ll Q^2$ is [24, 25]

$$d\sigma^{\mathrm{SIDIS}} \propto \Phi(x, \mathbf{k}_\perp) \otimes d\sigma^{eq} \otimes D(z, \mathbf{p}_\perp) \tag{1.36}$$

and in the case of Drell-Yan, under the same assumption that $\mathbf{k}_\perp$ of both partons is much smaller than the hard scale, the invariant mass of the lepton pair, [12, 26]

$$d\sigma^{\mathrm{DY}} \propto \Phi(x_1, \mathbf{k}_\perp^{\mathbf{1}}) \otimes \Phi(x_2, \mathbf{k}_\perp^{\mathbf{2}}) \otimes d\sigma^{q\bar{q}\to ll'} \tag{1.37}$$

where the longitudinal momentum fractions of the partons are fixed by external kinematics. On the contrary, TMD-factorization has been shown to break down in hadron-hadron collisions [27]. Thus, in the processes relevant for the measurement of Eq. (1.35), i.e. SIDIS and DY, factorization holds.

Non-zero quark Sivers TMD functions were reported, almost simultaneously, by the HERMES [28], COMPASS [29] and JLAB [30] experiments in semi-inclusive deep-inelastic scattering off transversely polarized nucleons. The milestone of measuring the sign flip of the Sivers function in Drell-Yan and SIDIS was recently achieved by the COMPASS facility [31]. It is fair to mention that the first experimental hint on this effect was reported by the STAR Collaboration using W and Z boson production from polarized p+p collisions at RHIC [32].

The COMPASS experiment is located at CERN and it is a fixed-target facility. In the case of Drell-Yan they bombard with a 190 GeV π^--beam, from the SPS accelerator, a transversely polarized ammonia (NH_3) target and detect muon-pairs. For SIDIS, the same target is shot with a μ^+-beam of 160 GeV. The hard scales explored lie in the range $1 < Q < 9$ GeV. The prominent feature of this experiment is its capability of performing DY and SIDIS measurements using mostly the same setup and in a similar kinematic region. Thus, there is no need to evolve the TMDs and the uncertainties are clearly reduced. The chosen observable sensitive to the sign flip is called *Sivers asymmetry* and measures the $\sin\phi_s$ (as defined in Fig. 1.10) modulation in the distribution of the produced hadrons or lepton pairs in SIDIS and DY respectively. The results of this observable in the SIDIS scenario are shown in Fig. 1.13. In this case, it only confirmed the non-zero value of f. However, when a combined analysis of the data from SIDIS and DY was performed the agreement with the change of sign hypothesis was successfully confirmed as shown in the right side of Fig. 1.13. The measurement of the Sivers asymmetry in a vastly extended kinematic region, as shown in Fig. 1.14, constitutes one of the major challenges in the physics program of the upcoming Electron Ion Collider [21].

1.2.3 GPDs and DVCS

Analogously to the TMDs case that sketch out a 3-D picture of the parton in momentum space, we would like to construct a function that encodes information on the distribution of partons both in coordinate space and longitudinal momentum. For that purpose, one has to integrate out the $\mathbf{k}_\perp$-dependence of Eq. (1.16) i.e.

$$G^q(x, \xi, \boldsymbol{\Delta}_\perp) = \int \frac{\mathrm{d}\eta^-}{2\pi} e^{-\mathrm{i}x\overline{P}^+\eta^-} \langle P'|\overline{\Psi}(0, \eta^-, \mathbf{0}_\perp)\gamma^+ G_{[\eta^-,0]}\Psi(0, 0, \mathbf{0}_\perp)|P\rangle\Big|_{\eta^+=\eta_\perp=0} \tag{1.38}$$

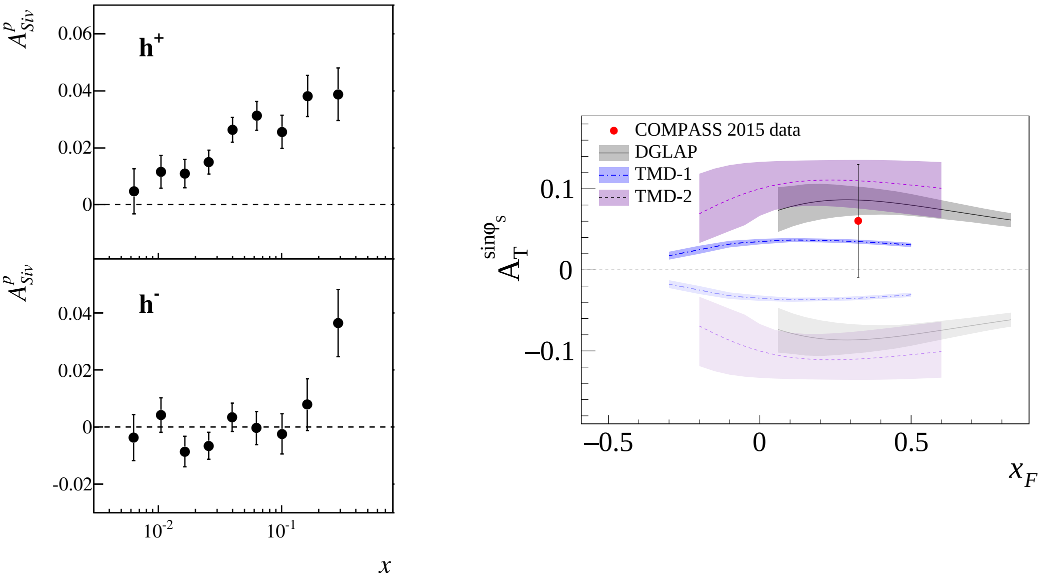

Figure 1.13: Left: Sivers asymmetry as a function of x for positive (top) and negative (bottom) hadrons as measured in SIDIS. Figure extracted from [33]. Right: The measured Sivers asymmetry as a function of x compared to theoretical predictions represented by dark-shaded (light-shaded) regions and evaluated with (without) the sign-change hypothesis. Figure extracted from [31].

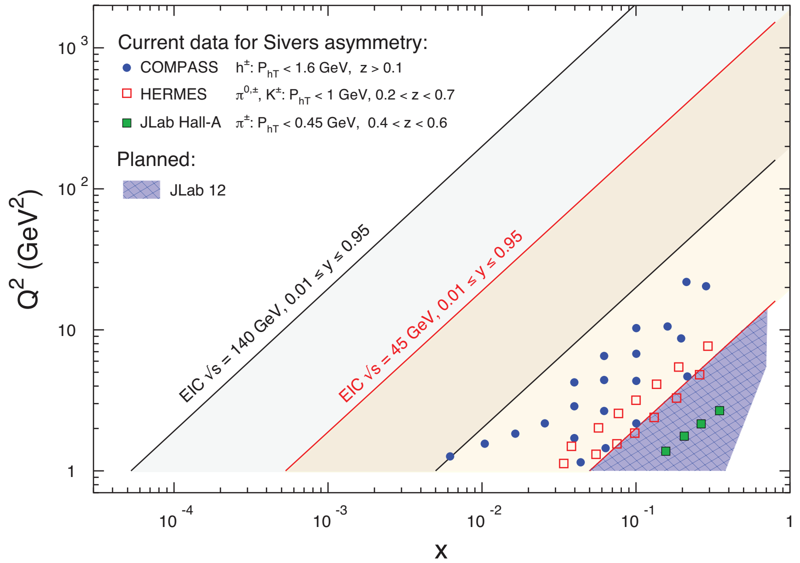

Figure 1.14: Kinematic coverage in the (x,Q^2)-plane expected in the Electron Ion Collider compared to other existing experiments for Sivers asymmetry. Figure extracted from [21].

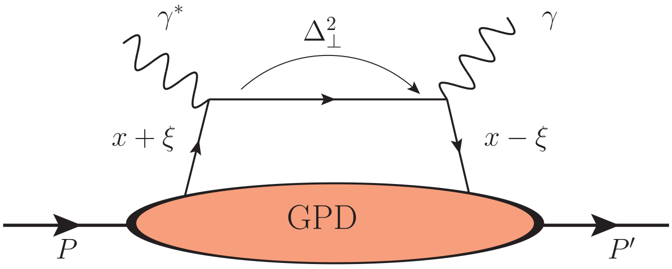

Figure 1.15: Kinematics of Deeply-virtual Compton Scattering.

where

$$\begin{aligned} \overline{P} &= (P+P')/2 \\ \xi &= \frac{\Delta^+}{\overline{P}^+} \\ x &= \frac{k^+}{P^+} \end{aligned} \tag{1.39}$$

and ξ, the light-cone momentum transferred to the proton, is called *skewness*. The function $G^q(x,\xi,\mathbf{\Delta}_\perp)$ is the definition of a Generalized Parton Distribution. There are several things of Eq. (1.38) that should be clarified. First, proton and parton momenta aren't the same before and after the scattering, consequently GPDs cannot be interpreted as a probability because they are defined from the overlap between different states. The transverse momentum component of the momentum transferred, $\mathbf{\Delta}_\perp$, is Fourier conjugate of the impact parameter, $\mathbf{b}_\perp$. Thus, the transverse structure of the target can be accessed thanks to $\mathbf{\Delta}_\perp$. In particular, when $\xi=0$ and the Fourier transform is performed, GPDs reduce to Impact Parameter Distributions defined as

$$I(x,\mathbf{b}_\perp) = \int \frac{\mathrm{d}^2\Delta_\perp}{(2\pi)^2} e^{-\mathrm{i}\mathbf{b}_\perp\cdot\mathbf{\Delta}_\perp} G(x,0,\mathbf{\Delta}_\perp) \tag{1.40}$$

as sketched in Fig. 1.3. $I(x,\mathbf{b}_\perp)$ does have a probabilistic interpretation.

The shape of the gauge link in Eq. (1.38) is identical to the PDFs case and therefore can be set to unity in the light-cone gauge. Finally, although the quark is assumed to be unpolarized, $\Gamma=\gamma^+$, the scattering may switch the helicity of the proton. Then, Eq. (1.38) can be rewritten as a sum of two terms

$$G^q(x,\xi,\mathbf{\Delta}_\perp) = \frac{1}{2\overline{P}^+}\left[H^q(x,\xi,\mathbf{\Delta}_\perp)\overline{u}(p')\gamma^+u(p) + E^q(x,\xi,\mathbf{\Delta}_\perp)\overline{u}(p')\frac{\mathrm{i}[\gamma^\alpha,\gamma^i]\Delta_\alpha}{2m_p}u(p)\right] \tag{1.41}$$

where $H^q(x,\xi,\mathbf{\Delta}_\perp)$ accounts for the case in which the proton helicity is unchanged and $E^q(x,\xi,\mathbf{\Delta}_\perp)$ that reflects the situation in which the target helicity flips.

Deeply-virtual Compton Scattering (DVCS) is one of the simplest scattering processes to constrain GPDs. The kinematics of DVCS is represented in Fig. 1.15: a highly virtual photon hits a parton, exchanges momenta with it and appears on-shell in the final state. A distinction should be made between the two transfered momenta that appear in Fig. 1.15. Q is related to the precision with which partons are resolved. In turn, $\mathbf{\Delta}_\perp$ determines the available resolution to determine the position of these partons inside the hadron. In other words, using an imaging analogue, Q^2 characterizes the probe to obtain the image and $\mathbf{\Delta}_\perp$ is a property of the object that is being imaged. Solely, when $\frac{1}{Q^2} \ll \frac{1}{\mathbf{\Delta}_\perp}$, formal proofs of factorization exist [34]. The factorization formula for DVCS adopts the following shape

$$d\sigma^{\gamma^* P} \propto \int_\xi^1 \mathrm{d}x d\sigma^{\gamma^* q} \otimes G(x, \xi, \mathbf{\Delta}_\perp; Q^2) \tag{1.42}$$

where an additional dependence on the factorization scale (set to Q^2) has been included in the GPDs. Notice that in the previous factorization formula, the longitudinal momentum fraction, x, appears in the convolution between the GPD and the partonic cross section and, therefore, this information is washed out. Performing the deconvolution of Eq. (1.42) is, in general, a not well-defined mathematical problem and greatly complicates the extraction of GPDs from experimental data [35]. On the contrary, both ξ and $\mathbf{\Delta}_\perp$ are directly measurable quantities. The only way to independently explore the (x, ξ)-dependence of GPDs from experimental data is through their evolution with the resolution scale Q^2 while fixing ξ and $\mathbf{\Delta}_\perp$. From a theoretical point of view the evolution equation that modifies the GPDs x-dependence when changing the resolution is known and acts in a similar way to DGLAP [10] (see Eq. (1.29)) not affecting ξ and $\mathbf{\Delta}_\perp$. However, on the experimental side, reconstructing the x-dependence of GPDs by changing Q^2 requires a large and precise data set in a wide range of Q^2 and ξ and will be feasible in the future EIC era as shown in Fig. 1.16 [21].

1.2.4 Lattice QCD

To finish the discussion on the structure of the proton, the role of lattice QCD will be briefly presented. For a complete review on this topic the reader is referred to [36].

Besides the common operator formalism any quantum field theory can be formulated within the path integral framework, a generalization of the action principle in classical mechanics. In the case of QCD it allows to compute the expectation value of any operator, $\mathcal{O}$, written in terms of quark and gluon fields, as

$$\langle \mathcal{O} \rangle = \frac{1}{Z} \int [\mathcal{D}q][\mathcal{D}\overline{q}][\mathcal{D}A] e^{\mathrm{i} \int \mathrm{d}^4 x \mathcal{L}_{\mathrm{QCD}}} \mathcal{O} \tag{1.43}$$

where $\mathcal{D}$ denotes a functional integral over the temporally ordered fields and $\mathcal{L}_{\mathrm{QCD}}$ is given by Eq. (1.1). The analytic solution to the path integral given by Eq. (1.43) remains unknown, so in principle, numerical techniques must be applied. For that purpose, in order to avoid the oscillatory behavior of the exponential formulated in Minkowski's metric, a rotation to Euclidean space is performed in the following way

$$t \to \mathrm{i}t_{\mathrm{E}} \Rightarrow \mathrm{i} \int \mathrm{d}^4 x \mathcal{L}_{\mathrm{QCD}} \to - \int \mathrm{d}^4 x \mathcal{L}_{\mathrm{QCD}}. \tag{1.44}$$

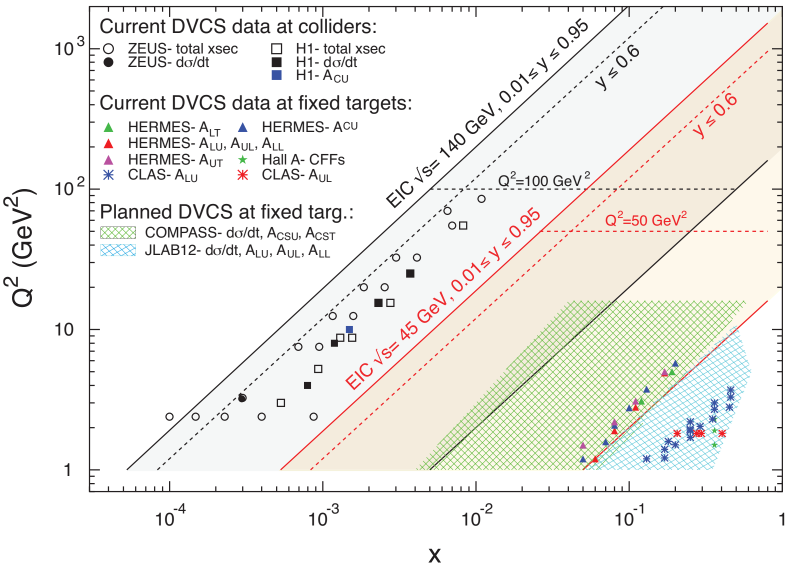

Figure 1.16: Kinematic coverage in the (x,Q^2)-plane expected in the Electron Ion Collider compared to other existing experiments for DVCS. Figure extracted from [21].

This formalism is the basis of lattice QCD that approximates the solution of Eq. (1.43) by discretizing the Euclidean space-time in form of a hypercubic lattice [37]. In Eq. (1.43) we have not imposed any constraint on the values of the coupling constant, it contains the full Lagrangian. The salient feature of lattice QCD is that it provides a first-principle approach to tackle the non-perturbative nature of QCD at small momentum transfer. However, when applying the method to the study of PDFs and GPDs a fundamental limitation appears: their x-dependence cannot be computed within lattice QCD due to the very Minkowski nature of their underlying operator structure. The definition of both PDFs and GPDs (see Eq. (1.17) and Eq. (1.38), respectively) contains bi-local operators on the light cone that introduce an intrinsic time-dependence. There is no continuous Lorentz transformation that rotates the non-local operator into Euclidean space or, in other words, converts the light-like separation into space-like, as in the case of TMDs [38][5]. Actually, a light-like path in Minkowski's space ($t^2-z^2=0$) shrinks into a single point in the Euclidean scenario ($t_E^2+z^2=0$) and so the ξ^--dependence is completely lost.

Although being far from a complete characterization of the hadron structure in terms of PDFs and GPDs within the lattice QCD framework, moments in the momentum fraction x given by

$$f^n = \int \mathrm{d}x x^{n-1} f(x) \tag{1.45}$$

can be computed in a controlled way by using the operator product expansion technique [40]. It provides an elegant solution to the non-locality of the operators in the PDFs and

---

[5] In the latter scenario, the complications come from the geometry of the gauge link and a full lattice calculation is not possible either up to date although progress is being made (see [39]).

GPDs definitions consisting in a Taylor expansion of the separation distance or, equivalently, converts the calculation into an infinite sum of local operators. The bi-local and local operators are then related through moments as defined by Eq. (1.45). In principle, computing all the moments of the functions would be equivalent to knowing their x-dependence. Nevertheless, only the lowest x moments up to $n=3$ have been obtained so far due to numerical limitations [41]. Therefore, a full reconstruction of the x functional form of PDFs and GPDs on the lattice remains unsettled.

1.3 Color Glass Condensate

The Color Glass Condensate (CGC), an effective field theory of high-energy QCD scattering, offers an ab-initio perspective on the characterization of hadron structure in a particular scenario: the small-x regime. For recent monographs on this topic see [42, 43].

The DIS experimental program carried out at HERA provided a picture of the hadronic wave function largely dominated by its gluonic component at small-x as it is shown in Fig. 1.9. This observation was expected from a theoretical point of view given that the probability of a quark to radiate gluons suffers a soft-divergence i.e. scales as $1/x$ (see Eq. (1.26)). However, the strong rise of the gluon PDF is not easy to reconcile with another experimental data set: the mild dependence of the total p+p cross section with the collision energy, $\sqrt{s}$, as presented in Fig. 1.17. Note that high-energy scattering is equivalent to small-x (see Eq. (A.6)). Typical values of x explored at RHIC and the LHC in central rapidities are: 10^{-2} and 10^{-4} respectively. Therefore, an indefinite growth of the gluon density would translate into a similar fast rise of the forward scattering amplitude and, eventually, lead to the violation of unitarity of the theory, a fundamental property of any QFT. Thus, it is clear that a novel dynamical mechanism is needed in order to tame the exponential behavior of the gluon PDF and describe the observed growth of the total p+p cross section.

Intuitively, when the proton becomes a densely populated system of gluons the assumption that interactions among them can be neglected in processes such as DIS breaks down. At a certain scale, $Q_s^2(x)$, the probability of two gluons to recombine into a single one is of the same order as the bremsstrahlung probability. $Q_s^2(x)$ is the so-called saturation scale and its value can be parametrically estimated using a very simple geometrical argument. The probability of two gluons to interact, κ, can be written as the product of the transverse gluon density times the typical gluon-gluon cross section i.e.

$$\begin{aligned} \kappa &= \rho^g \cdot \sigma^{gg\to q} \\ \kappa &\sim \frac{xf^g(x,Q^2)}{\pi R_h^2} \cdot \frac{\alpha_s}{Q^2} \end{aligned} \tag{1.46}$$

where R_h is the radius of the hadron and f^g represents the gluon PDF. κ is known as *packing factor* and it should be of order 1 when evaluated at $Q^2=Q_s^2$, thus

$$Q_s^2(x) \sim \frac{\alpha_s}{\pi R_h^2} xf^g(x,Q^2). \tag{1.47}$$

In the nuclear case, $R_A \sim A^{1/3} R_h$. Therefore, saturation effects kick in at lower energies in nuclear collisions. In this paradigm it is assumed that $Q_s \gg \lambda$ i.e. it is a semi-hard scale.

Hence, the evolution equations should be modified accordingly to include not only radiative terms but gluon fusion processes as well. Using a very schematic notation, DGLAP evolution can be expressed as

$$\frac{\partial f(x,Q^2)}{\partial \ln Q^2} \propto \mathcal{P} \otimes f(x,Q^2) \tag{1.48}$$

where $\mathcal{P}$ accounts for the splitting function given in Eq. (1.29). The solution of this linear differential equation has an exponential shape. In turn, when saturation effects are taken into account the x-evolution of the gluon Weiszacker-William distribution, $\phi(x, \mathbf{k}_\perp)$, also known as unintegrated gluon distribution (uGD)[6], is given by the so-called Balitsky-Kovchegov equation (BK) [45, 46]

$$\frac{\partial \phi(x,\mathbf{k}_\perp)}{\partial \ln(1/x)} \propto \mathcal{K} \otimes \phi(x,\mathbf{k}_\perp) - \phi(x,\mathbf{k}_\perp)^2 \tag{1.49}$$

where $\mathcal{K}$ is the evolution kernel that resums terms $\ln 1/x$ to all orders including diagrams such as the one depicted in Fig. 1.18. The non-linear nature of Eq. (1.49) results into a softening of the endless growth of $\phi(x, \mathbf{k}_\perp)$ and its integrated version at small-x as expected.

The consideration of non-linear evolution equations is one of the bedrocks of the Color Glass Condensate. However, this is not the end of the story. We can benefit from the fact that a hadron is, effectively, a collection of gluons at high energies to justify the applicability of semi-classical techniques to compute, from first principles, the functional form of $\phi(x, \mathbf{k}_\perp)$.

First of all, the condition $Q_s \gg \lambda$ defines a weak coupling regime where $\alpha_s(Q_s^2) \ll 1$. Further, the system is characterized by large occupation numbers that can be expressed in terms of creation and annihilation operators as

$$N = a^\dagger a \gg 1. \tag{1.50}$$

Therefore, when compared to the Heisenberg commutator that quantifies the quantum fluctuations

$$N \gg [a^\dagger, a] = 1. \tag{1.51}$$

The condition given by Eq. (1.51) defines a classical system. Then, the strong gluon fields characterizing the hadronic wave function at small-x can be treated as classical gauge fields A^a_μ. The origin for *color* and *condensate* in the name of the theory is now clear: we are dealing with a largely populated system of colored particles. To understand

---

[6] It should be noted that the conventional gauge-invariant operator definition of a gluon TMD differs from the Weiszacker-William distribution valid in the $x \ll 1$ regime. A study on how to relate both regimes was presented in [44]. Further, the unintegrated gluon distribution is related to the usual PDF by $f^g(x,Q^2) = \int^{Q^2} \mathrm{d}^2 k_\perp \phi(x, \mathbf{k}_\perp)$.

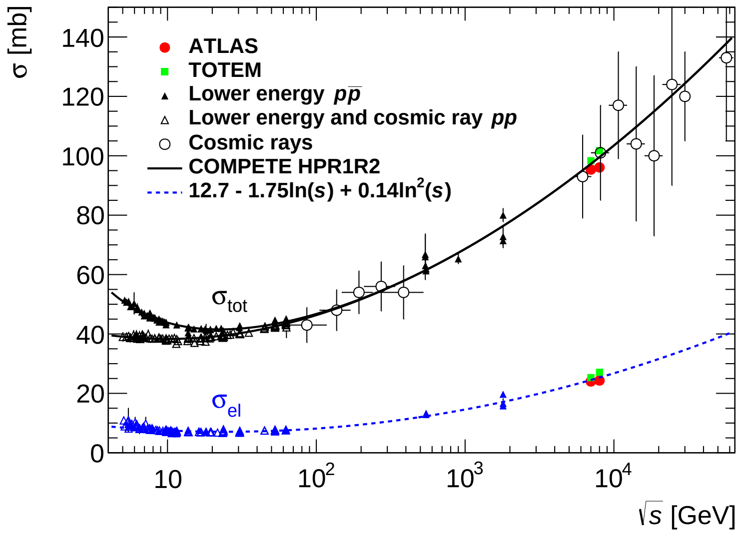

Figure 1.17: Total and elastic cross section measurements as a function of the centre of mass energy. Figure extracted from [47].

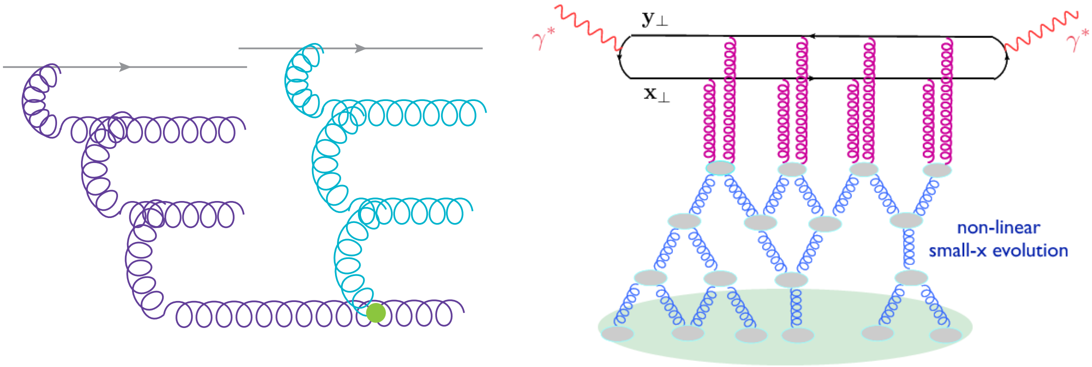

Figure 1.18: Left: Recombination of gluons coming from two different parton cascades. Right: Sketch of an electron-nucleus collision in the dipole model including multiple scatterings and non-linear small-x evolution. Figure extracted from [48].

where does the *glass* comes from we have to take a look into the space-time scales of the problem. In the infinite momentum frame, the hadron has a large light-cone momentum $P^\mu = (P^+, 0, \mathbf{0}_\perp)$. Partons with a large momentum fraction of its parent hadron, $p^+ = xP^+$, are, due to Heisenberg's uncertainty principle, sharply localized on the light-cone $\Delta x^- \sim 1/p^+$. Further, the light-cone temporal extent of these fluctuations, $\Delta x^+ \sim p^+/k_\perp$, is long. On the contrary, by using the same arguments, the small-x partons are short-lived and delocalized over larger distances. Therefore, a clear separation between *fast* and *soft* modes can be made in terms of their momentum fraction: large for the former and small for the latter. The definition of a glass is *a disordered system which evolves very slowly relative to natural time scales* [42]. In the same way, the soft modes, short-lived, see the fast ones as an infinitely thin and frozen configuration of color charges. More concretely, the valence partons, that carry a large momentum fraction, are identified with the fast modes and the small-x gluons with the soft ones. Thus, small-x gluons are described by classical fields A^a_μ. The fast partons propagate in an eikonal fashion so that they do not deviate from their light-cone trajectory $x^+ = 0$ (see Appendix C). To compute physical observables they are integrated out as they are not considered dynamical fields but a source of a color current in the $+$ -direction

$$J^a_\mu(x) = \delta^+_\mu \delta(x^-) \rho^a(x^-, \mathbf{x}_\perp) \tag{1.52}$$

where $\rho^a(x^-, \mathbf{x}_\perp)$ defines the density of color charges and it is independent of x^+ as we consider that the color charges are frozen. It is treated as a stochastic variable described by a weight functional $W_\Lambda[\rho]$. Finally, the relation between A^a_μ and J^a_μ is given by Yang-Mills equations of motion

$$[D_\mu, F^{\mu,\nu}_a] = J^\nu_a \tag{1.53}$$

that is, the analog of Maxwell equations in the non-Abelian case. The current J^ν must be covariantly conserved

$$[D_\nu, J^\nu] = 0 \Rightarrow \partial_\nu J^\nu = ig[A_\nu, J^\nu]. \tag{1.54}$$

This equation clearly shows that the produced gauge field has a feedback on the current itself. Putting together Eqs. (1.53) and (1.54) we end up with a system of equations that can be solved iteratively expanding in powers of ρ. The exact solution for the gauge field in the covariant gauge ($\partial^\mu A_\mu = 0$) is

$$A^{\mu,a} = \delta^{\mu +} b^a(x) \;\text{ with }\; -\nabla^2_\perp b^a(x) = \rho^a(x). \tag{1.55}$$

Thus, to compute any observable that depends on the small-x d.o.f in the CGC formalism one has to solve Eq. (1.53) first and afterwards average over all possible color configurations

$$\langle \mathcal{O}[A_\mu] \rangle = \int [d\rho] W_\Lambda[\rho] \mathcal{O}[A^\mu[\rho]]. \tag{1.56}$$

Now we have all the elements to fulfill the original purpose of this Section: compute $\phi(x, \mathbf{k}_\perp)$ within the CGC. Its definition is given by

$$\phi(x, \mathbf{k}_\perp) = \int \frac{d^2 r_\perp}{2\pi r^2_\perp} e^{-i\mathbf{r}_\perp \cdot \mathbf{k}_\perp} \left(1 - S(x, \mathbf{r}_\perp; A)\right) \tag{1.57}$$

where

$$S(x,\mathbf{r}_\perp;A)=\frac{1}{N_\mathrm{c}}\int[d\rho]W_\Lambda[\rho]\mathrm{Tr}\left[\mathcal{U}(\mathbf{x}_\perp)\mathcal{U}^\dagger(\mathbf{y}_\perp)\right] \tag{1.58}$$

represents the propagation of a $q\bar{q}$-dipole, originated from a photon, through the background color field of the target and therefore, is expressed in terms of Wilson lines. The process is sketched in the right pannel of Fig. 1.18. The missing ingredient is a specific functional form for $W_\Lambda[\rho]$, a model-dependent object. In the McLerran-Venugopalan model [49] a Gaussian probability distribution for the color charges is assumed

$$W_{\mathrm{MV}}[\rho]\propto\exp\left[-\int\mathrm{d}x^-\mathrm{d}^2x_\perp\frac{\rho^a(x^-,\mathbf{x}_\perp)\rho^a(x^-,\mathbf{x}_\perp)}{2\mu^2(x^-)}\right] \tag{1.59}$$

where $\mu^2(x^-)$ is the per nucleon charge density and it is related to the saturation scale. This ansatz, justified by the central limit theorem when the number of color charges is large enough as in the nuclear case, greatly simplifies the calculations involved in Eq. (1.57). Further, it leads to trivial correlators between charges

$$\langle\rho^a(x^-,\mathbf{x}_\perp)\rho^b(y^-,\mathbf{y}_\perp)\rangle=\delta^{ab}\delta^2(\mathbf{x}_\perp-\mathbf{y}_\perp)\delta(x^--y^-)\mu_A^2. \tag{1.60}$$

The following parametric solution for $\phi(x,\mathbf{k}_\perp)$ is obtained [48]

$$\phi^{\mathrm{MV}}(x,\mathbf{k}_\perp)\sim\begin{cases}\ln k_\perp^2/Q_s^2 & \text{for } k_\perp\ll Q_s\\ Q_s^2/k_\perp^2 & \text{for } k_\perp\gg Q_s\end{cases}. \tag{1.61}$$

Remarkably, the unintegrated gluon distribution saturates for $k_\perp\ll Q_s$ and exhibits a power-like behavior for large transverse momenta as in the usual perturbative calculation.

Up to now we have performed a purely classical calculation. It is well suited to describe gluon modes with a value of x close to that at which the classical action is defined, $x\sim\Lambda/P^+$. However, when trying to describe modes with momentum fraction $x'\sim\Lambda'/P^+$ much smaller than x quantum corrections are large. This is due to the fact that, as in the case of the factorization scale (see discussion in Sec. 1.2.1), the momentum scale Λ at which the separation between soft and fast modes is done is totally arbitrary and the interactions do not disappear as we move away from this scale to. These corrections can be resummed via a renormalization group equation, JIMWLK, for the weight functional of color charges that adopts the following Hamiltonian form [50]

$$\frac{\partial W_Y[\rho]}{\partial Y}=HW_Y[\rho] \tag{1.62}$$

where $Y=\Lambda/\Lambda'$ and H includes the kernel of the evolution that is a non-linear function of ρ. Thus, the purely classical calculation can be adapted to contain the quantum modifications by replacing $W_\Lambda[\rho]$ with $W_Y[\rho]$, a modified statistical weight, whose variation with the scale, Y, is given by the JIMWLK equation. Then, the McLerran-Venugopalan model is used as an initial condition for this evolution equation. In practice, solving the JIMWLK equation is highly non-trivial so, usually, its mean field and large-N$_c$ approximation, i.e.

the BK equation schematically given in Eq. (1.49), has become the starting point for phenomenological studies. At this point we have learned how the nonlinear evolution leads to a forward scattering amplitude that doesn't violate unitarity i.e. it is always smaller than one. Nevertheless, as will be discussed in Sec. 2.3.1, the total proton-proton cross section obtained from it rises as a power of energy violating the Froissart-Martin bound that states that hadronic total cross sections in the asymptotic limit, $s \to \infty$, should not grow faster than $\propto \ln^2 s$ [51]. This condition arises from very general properties of the scattering such as unitarity, crossing or analiticity. The violation of the Froissart bound has been confirmed both theoretically [52] and by numerically solving BK evolution including impact parameter dependence [53]. This is not a flaw exclusive of BK evolution but of any perturbative framework that is applied beyond its limit of validity. Therefore, CGC techniques should be complemented with non-perturbative effects to make the total cross section satisfy the Froissart bound.

The methodology to compare CGC and data is the same as in the PDF scenario: the unintegrated gluon distributions are obtained by fits to HERA data and later on applied in other physical processes sensitive to the small-x regime provided that the corresponding factorization theorems apply. In Fig. 1.19, two particular cases in which CGC calculations successfully met experimental data are depicted. The unintegrated gluon distributions constrained by HERA data for the proton were scaled to the nuclear case and used to compute the multiplicity of charged particles in a heavy ion collision at a centre of mass energy per nucleon pair, $\sqrt{s_{NN}}$, of 2.76 TeV. The agreement between theory and data is, in both cases, excellent and confirms the predictive power of the Color Glass Condensate. This phenomenological success reinforces the status of the CGC as the best candidate to approximate QCD dynamics in high-energy scattering or, equivalently, the small-x regime.

1.4 This thesis in the context of the proton structure

All in all, significant progress has been made until reaching our current understanding of proton structure despite its non-perturbative nature. Briefly, our knowledge of proton structure consists of:

- **HERA legacy**: The longitudinal momentum information of the partons is well constrained thanks to a remarkable synergy between theoretical developments, such as factorization theorems and renormalization group equations, and experimental effort in terms of DIS data among others. Currently, PDF sets at NNLO accuracy are available for high-precision LHC phenomenology studies with uncertainties at the percent level in some kinematic window. Further, DIS at HERA opened the door to study the small-x regime. An excellent agreement between data and experiment has been achieved both in the framework of BFKL dynamics and saturation physics (see left pannel in Fig. 1.19) suggesting the necessity to go beyond DGLAP by including non-linear terms in the evolution equations in this kinematic regime.

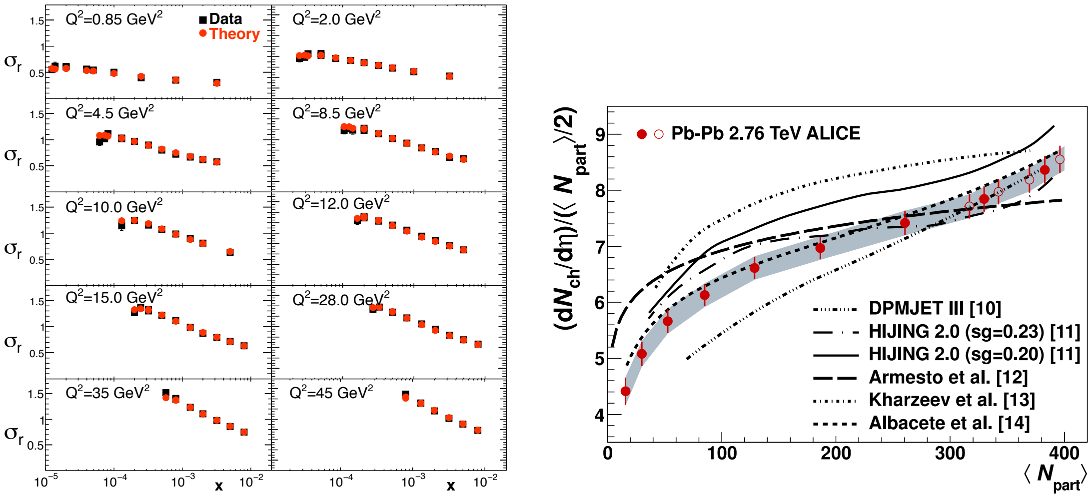

Figure 1.19: Left: Comparison of the CGC calculation of the reduced cross section, a combination of structure functions, with HERA data on electron-proton collisions. Figure extracted from [54]. Right: Measured multiplicities in a lead-lead collision at LHC energies compared to different theoretical predictions. Figure extracted from [55].

- **3-D imaging**: Besides its intrinsic theoretical interest, the necessity of going beyond PDFs stems from the study of exclusive observables in which the final state of the interactions is partially or fully measured. In contrast with DIS where it is integrated out and the only relevant scale is x. Several complications appear when trying to extract the 3-D image of a proton in momentum or coordinate space:
  - **TMDs**: One of the recent breakthroughs of the field is the experimental confirmation by the COMPASS Collaboration of the Sivers asymmetry, a robust prediction of the theory that emerges from the gauge-invariance properties of QCD. This measurement evidences their non-universality i.e. they are process dependent. Further, a timely line of research consists in exploring the small-x limit of TMDs within the CGC approach [56].
  - **GPDs**: The mixed information on longitudinal momentum, intrinsic (x) and transfer (ξ), and transverse coordinates described by GPDs encounters a major drawback: the explicit dependence on the intrinsic longitudinal momentum of partons is lost in experimental measurements such as DVCS. This is so because it always appears convoluted with a function representing the hard-scattering process. The information on the separate dependence can only be obtained by measuring the Q^2 variation of exclusive processes at a given ξ.

  Apart from the theoretical developments, the wide kinematic coverage that will be explored by the upcoming Electron Ion Collider woud allow to perform tomographic images of the nucleon with unprecedented precision.

- **Ab-initio methods**: Even though hadron structure is dominated by long-range phenomena, the situation concerning its theoretical description is not completely hopeless. Two well-defined approaches tackle this question by means of first principles calculations:

- **Lattice QCD methods**: Being perhaps the best approximation to QCD, this technique is not able to describe the full longitudinal momentum dependence by construction i.e. the action is discretized in Euclidean space. However, moments of these distributions are formally well defined on the lattice and the main reason behind the fact that only low moments have been computed so far are numerical bottlenecks.
- **Color Glass Condensate**: The experimental confirmation of the strong growth of gluonic densities at small values of x was the smoking gun for the construction of this effective field theory of QCD in the small-x regime. The large gluonic densities enable a semi-classical description of the gluon content in hadronic wavefunctions. Deep-inelastic scattering is the perfect playground to test the CGC. Further, nuclei can be initially characterized within the CGC in ultra high energy heavy ion collisions. Its role in the description of QGP physics will be explored in detail in Chapter 4.

In this thesis a phenomenological study of the spatial distribution of subnucleonic degrees of freedom in the proton is performed. We base our model on features of elastic scattering data of proton-proton interactions at low ($\sqrt{s} = 62.5$ GeV) and high ($\sqrt{s} = 7$ TeV) energies as measured in ISR and LHC respectively. The momentum information is omitted and thus, a natural extension of our studies would be to investigate the x-dependence of the considered degrees of freedom in our model and construct GPDs out of it. In the next Chapter a detailed presentation of the ingredients of our description together with the experimental data set is provided.

Into the hollow;
Queens of the Stone Age.

2

The *hollowness* effect

The best way to constrain the transverse spatial structure of the proton is to scrutinize the available elastic scattering data. The analysis of the elastic proton-proton differential cross-section at collision energy $\sqrt{s}=7$ TeV measured by the TOTEM Collaboration [57] has revealed a new, intriguing feature of hadronic interactions: at high energies, the inelasticity density of the collision does not reach a maximum at zero impact parameter. This phenomenon, not observed before at lower collision energies, has been referred to as *hollowness* effect [58].

In this Chapter, we present a dynamical explanation of the *hollowness* effect. Our analysis is based on three main ingredients: we rely on gluonic *hot* spots inside the proton as effective degrees of freedom for the description of the scattering process. Next, we assume that some non-trivial correlation between the transverse positions of the hot spots inside the proton exists. Finally, we build the scattering amplitude from a multiple scattering, Glauber-like series of collisions between hot spots. In our approach, the onset of the *hollowness* effect is naturally explained as due to the diffusion or growth of the hot spots in the transverse plane with increasing collision energy. This microscopic realization of the hollowness effect sets stringent constraints on the spatial distribution of the gluonic hot spots.

This Chapter is based on the publication:

- J. L. Albacete, A. Soto-Ontoso, Hot spots and the hollowness of proton–proton interactions at high energies, Phys. Lett. B770 (2017) 149–153 [59].

2.1 Proton-proton elastic scattering

Counterintuitively, one third of the times two protons collide at LHC energies they do not produce new charged particles but exchange momentum while retaining their identity after the scattering i.e. they interact elastically. As depicted in Fig. 1.3[1], elastic scattering provides a sensitive probe of the proton transverse structure with a resolution given by the momentum transfer $\vec{q}$, Fourier conjugate to the impact parameter $\vec{b}$. The impact parameter is defined as the distance between the centres of the two protons. The larger the momentum transfer, the deeper one looks inside the proton. It should be noted that at very small-$|t|$ or, equivalently, when the separation between the protons is large, the electromagnetic force becomes dominant. Then, the elastic interaction is described by Coulomb scattering and reaches the same magnitude as the strong interaction when $t_0 \sim 1/(14\sigma_{\rm tot}) \sim 3\times10^{-4}$ GeV^2 at $\sqrt{s}=7$ TeV [60]. In our analysis, we will work in the nuclear region, i.e. we consider $t \geq 5\times10^{-3}$ GeV^2, such that both the electromagnetic and the Coulomb-nuclear interference terms will be neglected. The differential elastic cross section is given by

$$\frac{{\rm d}\sigma_{\rm el}}{{\rm d}t} = \frac{1}{4\pi}|T_{\rm el}(s,t)|^2 \tag{2.1}$$

where $T_{\rm el}(s,t)$ is the elastic scattering amplitude that depends both on the absolute value of the momentum transfer $|t| = q^2$, and the energy of the collision s. In the following we will work with the Fourier-Bessel transform of the elastic scattering amplitude computed as

$$\widetilde{T}_{\rm el}(s,\vec{b}) = \int \frac{d^2q}{2\pi} e^{{\rm i}\vec{q}\cdot\vec{b}} T_{\rm el}(s,q^2) = \int qJ_0(qb)T_{\rm el}(s,q^2)dq \tag{2.2}$$

with J_0 the zeroth Bessel function and azimuthal symmetry has been assumed in the last step given that both target and projectile are unpolarized. The optical theorem establishes the relation between the forward ($t=0$) scattering amplitude and the total cross section. With this normalization the total cross section, $\sigma_{\rm tot}$, the $2\to2$ elastic cross section, $\sigma_{\rm el}$, and the ratio of real and imaginary parts of the scattering amplitude, ρ, read, respectively

$$\sigma_{\rm tot}(s) = 2{\rm Im}T_{\rm el}(s,0) = 2\int {\rm d}^2b\ {\rm Im}\widetilde{T}_{\rm el}(s,\vec{b}) \tag{2.3}$$

$$\sigma_{\rm el}(s) = \int {\rm d}^2b \left|\widetilde{T}_{\rm el}(s,\vec{b})\right|^2 \tag{2.4}$$

$$\rho(s,t) = \frac{{\rm Re}T_{\rm el}(s,t)}{{\rm Im}T_{\rm el}(s,t)}. \tag{2.5}$$

These two expressions can be combined into

$$\sigma^2_{\rm tot}(s) = \frac{16\pi}{1+\rho^2(s,0)}\frac{{\rm d}\sigma_{\rm el}}{{\rm d}t}\bigg|_{t=0} \tag{2.6}$$

therefore, the differential elastic scattering data can be exploited to determine the total cross section, as it is actually done in experiment. The value of ρ at $t=0$ is usually taken from an external source (such as COMPETE fits [61]) or it can also be extracted by analyzing the interference between Coulomb and hadronic contributions to ${\rm d}\sigma_{\rm el}/{\rm d}t$

---

[1] The notation *form factor* is subtle as it usually refers to the charge distribution as measured in electron-proton scattering. In our case, we parametrize the matter distribution.

[62]. In the latter scenario, the t-dependence of ρ is sensitive to the electromagnetic form factor of the proton and the impact of different parametrizations has been considered in the literature [63] being the subject to an on-going debate [64]. In any case, as it enters the computation of $\sigma_{\rm tot}^2$ like $1+\rho^2$ the influence of any errors in the determination of ρ is small.

From an experimental point of view, the measurement of elastic processes is extremely challenging. In order to benefit from Eq. (2.6), a high degree of efficiency in detecting small momentum transfers or, equivalently, protons with small deflection angles (see Eq. (A.5)) is mandatory to reduce the uncertainties related to the extrapolation of $\mathrm{d}\sigma_{\rm el}/\mathrm{d}t$ to $t=0$, given that this value cannot be measured directly. To be sensitive to small angles, dedicated detectors capable of measuring tracks of scattered protons flying at millimeter distance from the beam pipe and special beam optics, with reduced beam divergencies, are needed. In proton-proton collisions at CERN, early measurements made at the Intersecting Storage Rings (ISR) at energies lying in the $23.4<\sqrt{s}<62.5$ GeV range and the most recent results achieved by TOTEM, with energies more than 100 times higher, shared the same detection set up. It is composed by forward (see Eq. (A.3))[2] tracking telescopes to detect the charged particles coming from inelastic p+p collisions together with a system of *roman pots* located far (~ 100 m) from the interaction point but as close as possible (<1 mm) to the beam. A roman pot, named after the CERN/Rome group that introduced this experimental technique in the early 1970s at ISR, is a movable beam insertion equipped with tracking detectors. During beam injection, these secondary vacuum chambers remain in a safe position and approach close to the beam for measurements thanks to vacuum bellows. In the case of TOTEM there are 26 roman pots, placed symmetrically, housing silicon strips. Thus, the scattered back-to-back protons hit the detectors and by computing their transverse motion from the interaction point to the roman pots, the scattering angle is determined. Once the scattering angle is known the momentum transfer $|t|$ in the elastic scattering can be reconstructed by using the relation given in Eq. (A.5) [65].

It is worth to mention that $\mathrm{d}\sigma_{\rm el}/\mathrm{d}t$ has been measured in proton-antiproton interactions at the S$\bar{\rm p}$pS collider [66] at CERN and the Tevatron [67] at Fermilab between $\sqrt{s}=546$ GeV and 1.96 TeV. Nevertheless, our study is restricted to the p+p case. The measurements of the elastic proton-proton differential cross-section at ($\sqrt{s}=7$ TeV) by the TOTEM Collaboration [57] at the LHC and at the maximum ISR energy ($\sqrt{s}=62.5$ GeV) [68] are presented in Fig. 2.1. The data on $\mathrm{d}\sigma_{\rm el}/\mathrm{d}t$ exhibits a high degree of structure and consequently its description is challenging from a phenomenological point of view. First, we will focus on the generic characteristics of the curve and afterwards on the energy dependence of the observed features.

- The small-$|t|$ behavior (before the pronounced dip) follows an almost exponential shape

$$\frac{\mathrm{d}\sigma_{\rm el}}{\mathrm{d}t} \sim \left(\frac{\mathrm{d}\sigma_{\rm el}}{\mathrm{d}t}\right)\Big|_{t=0} e^{-B|t|} \tag{2.7}$$

---

[2] At TOTEM the tracking telescopes cover $3.1 \leq |\eta| \leq 6.5$.

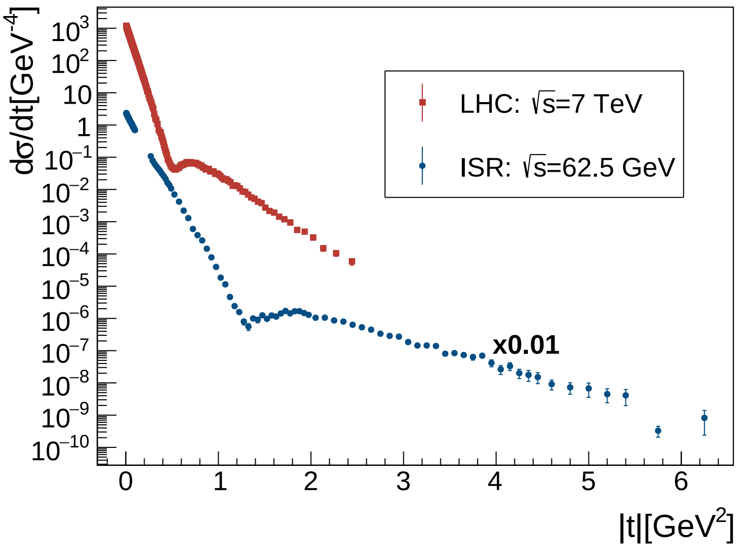

Figure 2.1: $\mathrm{d}\sigma_{\mathrm{el}}/\mathrm{d}t$ as a function of the momentum transfer at $\sqrt{s}=7$ TeV (red squares) and $\sqrt{s}=62.5$ GeV (blue dots). Note that the ISR data has been multiplied by 0.01 for readability. The error bars account for statistical uncertainties.

where $B \equiv B(s)$ is called the slope parameter. Hints of deviation from the single exponential behavior i.e. $B(s) \to B(s,t)$ were reported in [69]. Naively, Fourier transforming the exponential function in Eq. (2.7) a Gaussian in impact parameter space $\sim e^{-b^2/B^2}$ is obtained. This points to a direct relation between the slope parameter and the only length in the game i.e. the proton radius.

- After the exponential fall-off, moving to larger $|t|$ values a pronounced dip is encountered followed by a secondary maximum. This characteristic pattern is univocal of diffractive phenomena. To make this point clear, let's consider Eq. (2.1) in impact parameter representation

$$\frac{\mathrm{d}\sigma_{\mathrm{el}}}{\mathrm{d}t} = \frac{1}{4\pi}|T_{\mathrm{el}}(s,\vec{q})|^2 = \frac{1}{4\pi}\left|\int d^2q e^{-\mathrm{i}\vec{q}\cdot\vec{b}}\widetilde{T}_{\mathrm{el}}(s,\vec{b})\right|^2. \tag{2.8}$$

Now, if the scattering is off a totally absorptive target that is circular with radius R (a black disk) the elastic amplitude is given by

$$\widetilde{T}_{\mathrm{el}}(s,\vec{b}) = \mathrm{i}\Theta(R-b) \tag{2.9}$$

where Θ is the Heaviside function. Then, the elastic differential cross section reads

$$\frac{\mathrm{d}\sigma_{\mathrm{el}}}{\mathrm{d}t} = \pi R^2 \frac{J_1^2(\sqrt{|t|}R)}{|t|}. \tag{2.10}$$

The series of minima of Eq. (2.10) are the zeroes of the Bessel function and their position is uniquely determined by the inverse size of the target. In other words, the

	B [GeV^{-2}]	$\|\mathbf{t}_{\text{dip}}\|$ [GeV^2]	$\mathbf{n}$ in t^{-n}
ISR: $\sqrt{s}$=62.5 GeV	13.3 ± 0.2	1.26 ± 0.03	–x–
LHC: $\sqrt{s}$=7 TeV	23.6 ± 0.6	0.53 ± 0.01	7.8 ± 0.3

Table 2.1: Parameters characterizing the shape of $\mathrm{d}\sigma_{\text{el}}/\mathrm{d}t$ at ISR and LHC energies. Systematic and statistical uncertainties have been added in quadrature. The –x– symbol indicates that this value has not been measured.

position of the dip in t-space scales as $1/\sigma_{\text{tot}}$ in this geometric interpretation. We see that an analogy can be established between high-energy elastic scattering off a black disk and light diffraction. Unavoidably, the parametrization given by Eq. (2.10) fails when confronted with the experimental data presented in Fig. 2.1 as the spatial structure of the proton is much more involved than the one of a black disk.

- At very high $|t|$, the differential elastic cross section can be described by a power-law t^{-n}. In this region, the QCD coupling constant is small enough so that perturbative techniques are applicable and the partonic structure of the proton is resolved. In particular, it was shown in [70] that an energy independent $\mathrm{d}\sigma_{\text{el}}/\mathrm{d}t \propto t^{-8}$ arises naturally when describing the proton-proton scattering as the interaction of valence quarks via the exchange of three gluons in a colorless state with negative charge parity. The three gluons exchange is the easiest perturbative realization of an object, the so-called *odderon*, that appears in Regge field theory, a pre-QCD formalism that will not be covered in this tesis. For a detailed review and the connections between Regge language and QCD the reader is referred to [71].

The values of the slope parameter B, the position of the diffractive minimum $|t_{\text{dip}}|$ and the exponent n of the power-law fall-off at large-$|t|$ for ISR and LHC energies are presented in Table 2.1. Firstly, although the ISR Collaboration did not make a dedicated analysis of the large-$|t|$ behavior of the elastic scattering amplitude, the experimental data was shown to be compatible with the three gluon exchange interaction in [72]. The measured value of n at an energy hundred times larger is also compatible with the pQCD model and thus reinforces the validity of this approach in the partonic regime. On the contrary, the position of the diffractive minimum is strongly energy-dependent: it is shifted towards smaller values of $|t|$ when the centre of mass energy is increased. As it has been previously mentioned, $|t_{\text{dip}}|$ can be related, in a geometric picture, to the inverse of the total cross section or, equivalently to the inverse size of the hadron R^{-2}. This phenomenon is known in the literature as the *shrinkage of the forward elastic peak* and was already observed before the advent of the LHC in the low energy scan performed at ISR from 23.5 to 62.5 GeV [68]. This result is confirmed by the opposite $\sqrt{s}$-dependence of the slope parameter, that scales as R^2. Indeed, both experimental observations suggest a diffusion of the proton interaction radius that naturally explains the raise of the total p+p cross section when increasing the collision energy (see Fig. 1.17). Naively, the measured values of the slope parameter would correspond to $R{\sim}0.72$ and 0.96 fm at ISR and LHC energies, respectively. A natural question arises at this point: what is the underlying dynamical mechanism responsible for the swelling of R. In Sec. 2.4 we argue that the measured growth of the total proton-proton cross section can be accounted for by the diffusion of the proton constituents and not by the proton itself. Further, in a radiative description of proton subnucleonics degrees of freedom this growth emerges naturally.

In the next Section it is shown how to learn from the inelasticity component of the collision by the data on $\mathrm{d}\sigma_{\mathrm{el}}/\mathrm{d}t$.

2.2 Extraction of $G_{\mathrm{in}}(s,b)$ from data

The conservation of total probability, also known as unitarity condition, constitutes one of the main bedrocks of scattering theory together with Lorentz invariance. Succinctly, it can be written as

$$SS^{\dagger} = I \tag{2.11}$$

where S and I are the scattering and identity matrices, respectively. Possibly, a more transparent form of Eq. (2.11) is given in terms of cross sections

$$\sigma_{\mathrm{tot}} = \sigma_{\mathrm{in}} + \sigma_{\mathrm{el}} \tag{2.12}$$

where σ_{in} is the total inelastic cross section. For our discussion, it is more convenient to formulate the unitarity condition in impact parameter space by means of Eqs. (2.5)

$$G_{\mathrm{in}}(s,\vec{b}) \equiv \frac{\mathrm{d}^2\sigma_{\mathrm{in}}}{\mathrm{d}^2\vec{b}} = 2\mathrm{Im}\widetilde{T}_{\mathrm{el}}(s,\vec{b}) - |\widetilde{T}_{\mathrm{el}}(s,\vec{b})|^2. \tag{2.13}$$

The most reliable relation between elastic and inelastic processes is provided by Eq. (2.13). The inelasticity density (or profile), $G_{\mathrm{in}}(s,\vec{b})$, is the probability of having a collision where new particles are produced at a given energy $\sqrt{s}$ and impact parameter $\vec{b}$. Conforming to unitarity, the inelasticity density is bound like $0 < G_{\mathrm{in}}(s,\vec{b}) < 1$ and the probabilistic interpretation holds. However, impact parameters cannot be measured experimentally. Therefore, the inelasticity density is not a physical observable. Up to now, the only known method to explore the functional shape of $G_{\mathrm{in}}(s,\vec{b})$ is based on the measurements of $\mathrm{d}\sigma_{\mathrm{el}}/\mathrm{d}t$. An important comment is in order. From Eq. (2.13) it is neat that through an accurate description of data on $\mathrm{d}\sigma_{\mathrm{el}}/\mathrm{d}t$ one is sensitive only to the square of the absolute value of the amplitude, not to real and imaginary parts independently i.e. not to the phase. Thus, the information gained with $\mathrm{d}\sigma_{\mathrm{el}}/\mathrm{d}t$ should be complemented with the description of other features of p+p interaction which do depend on the phase such as the real to imaginary ratio of the scattering amplitude at $t = 0$. It would be desirable to obtain a complete description of $\rho(s,t)$, however, the experimental information is scarce and from a theoretical point of view, as it has been mentioned, involves the modeling of the electromagnetic structure of the proton, a source of uncertainties.

The procedure to construct $G_{\mathrm{in}}(s,\vec{b})$ starts by fitting the data on $\mathrm{d}\sigma_{\mathrm{el}}/\mathrm{d}t$ with a certain parametrization of the elastic scattering amplitude. The real and imaginary parts of $T_{\mathrm{el}}(s,t)$ should be also constrained by describing $\rho(s,0)$. Next, $T_{\mathrm{el}}(s,t)$ is Bessel-Fourier transformed to impact parameter space as expressed in Eq. (2.2). At this point,

	$\mathbf{a_1}$	$\mathbf{a_2}$	$\mathbf{b_1}$	$\mathbf{b_2}$	$\mathbf{b_3}$
ISR	-0.31 ± 0.02	30.22 ± 4.14	-0.99 ± 0.02	-8.41 ± 0.48	-4.44 ± 0.16
LHC	-7.92 ± 0.49	99.56 ± 4.56	-2.45 ± 0.04	-11.00 ± 0.17	-6.05 ± 0.27

	$\mathbf{d_1}$	$\sigma_{\rm tot}$ [mb]	ρ
ISR	-4.24 ± 0.11	43.3	0.1
LHC	-5.98 ± 1.00	98.3	0.12

Table 2.2: ISR and LHC parameters, expressed in [GeV^{-2}], extracted from the fit with the parametrization given by Eq. (2.14) and the corresponding $\chi^2/{\rm d.o.f.}$ $\chi^2/{\rm d.o.f}$ is computed taking into account the data statistical uncertainties. The values of a_3 and c_1 are fixed to reproduce the measured values of $\sigma_{\rm tot}$ and ρ at each energy taking into account the uncertainties.

one has to make sure that the t-range measured experimentally is sufficient to carry out the numerical Fourier transformation with a satisfactory accuracy. In the case of TOTEM data, this issue was tackled in [73] where it was explicitly shown that the integrand of Eq. (2.2) vanishes long before reaching the maximum measured value of t. Finally, the inelasticity density can be constructed from $\widetilde{T}_{\rm el}(s,\vec{b})$ by means of Eq. (2.13). Albeit several analysis of ISR and LHC data on ${\rm d}\sigma_{\rm el}/{\rm d}t$ have been recently presented in the literature [73, 74, 75, 76], resulting into similar quality fits, we have performed our own independent analysis whose details are presented hereafter.

To describe the data on ${\rm d}\sigma_{\rm el}/{\rm d}t = (1/4\pi)\,|T_{\rm el}(s,t)|^2$ we resort to the following generic parametrization:

$$\begin{aligned}
{\rm Im}T_{\rm el}(s,t) &= a_1 e^{b_1 t} + a_2 e^{b_2 t} + a_3 e^{b_3 t}\,, \\
{\rm Re}T_{\rm el}(s,t) &= c_1 e^{d_1 t}\,,
\end{aligned} \tag{2.14}$$

with the fit parameters (a_i, b_i, c_i, d_i) being energy dependent and (a_i, c_i) subjected to the constraints:

$$\sigma_{\rm tot} \;=\; 2\sum_{i=1}^{3} a_i = \begin{cases} 43.32 \pm 0.34~{\rm mb} & \text{ISR} \\ 98.3^{+2.8}_{-2.7}~{\rm mb} & \text{LHC} \end{cases} \tag{2.15}$$

$$\rho \;=\; c_1 \sum_{i=1}^{3} 1/a_i = \begin{cases} 0.095 \pm 0.018 & \text{ISR} \\ 0.14^{+0.01}_{-0.08} & \text{LHC} \end{cases} \tag{2.16}$$

where the values of $\sigma_{\rm tot}$ and ρ correspond to the experimental measurements given in [57, 68]. For the LHC case we use the extrapolated ρ value provided by the COMPETE Collaboration [61], same as the TOTEM collaboration in their data analysis. With this set up we obtain a very good description of ${\rm d}\sigma_{\rm el}/{\rm d}t$, $\chi^2/{\rm d.o.f} = 1.1$ for ISR and $\chi^2/{\rm d.o.f} = 2$ for the LHC as depicted in Fig. 2.2. The minimization procedure has been performed with the MINUIT package [77]. In Table 2.2 the numerical values of the fit parameters are presented. Nevertheless, their physical interpretation is unclear as Eq. (2.14) is just

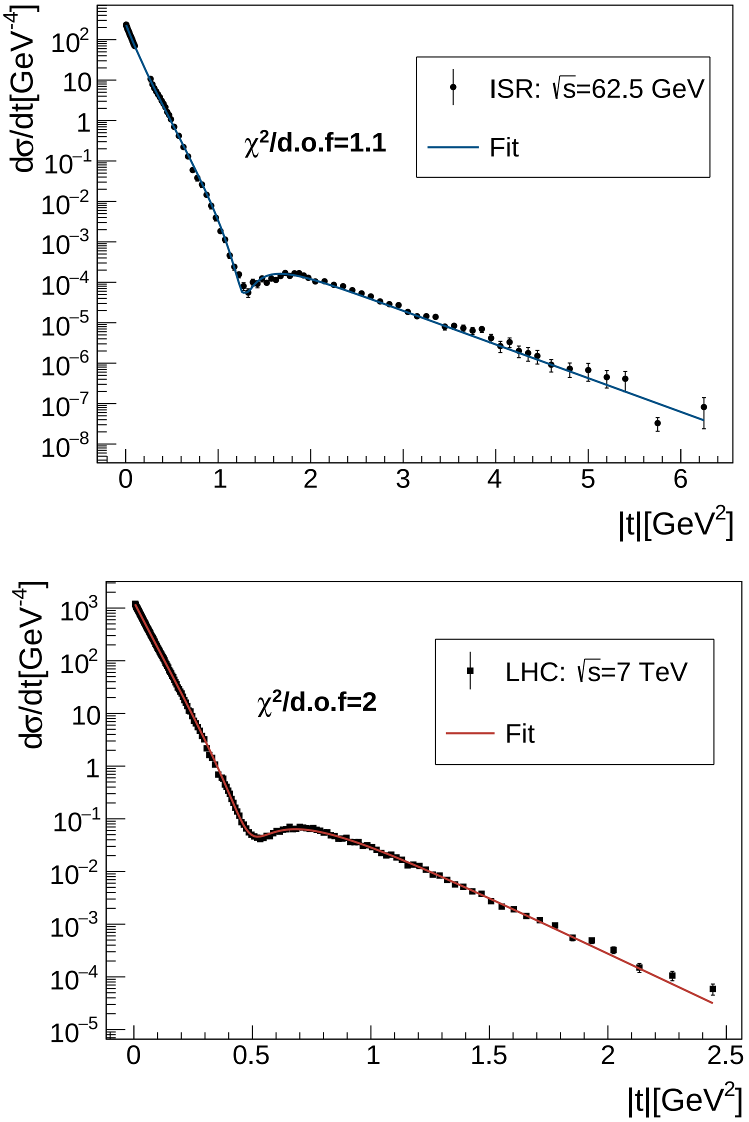

Figure 2.2: Fit to $\mathrm{d}\sigma_{\mathrm{el}}/\mathrm{d}t$ with the parametrization given in Eq. (2.14) for $\sqrt{s} = 62.5$ GeV (top) and $\sqrt{s} = 7$ TeV (bottom). The error bars account for statistical uncertainties. $\chi^2(\alpha) = \sum_{i=1}^{n}(f(x_i, \alpha) - e_i)^2/\sigma_i^2$ where f is the theoretical parameterization that depends on the vector of free parameters being fitted α and e_i are the individual measurements with uncertainties σ_i.

a general parametrization of the elastic scattering amplitude flexible enough as to allow a very good description of data. Fits with a larger number of free parameters lead to similar values of $\chi^2/\mathrm{d.o.f.}$ Thus, in consistency with other works in the literature, we find that any fit to the highly precise and structured data on $\mathrm{d}\sigma_{\mathrm{el}}/\mathrm{d}t$ with $\chi^2/\mathrm{d.o.f} \sim \mathcal{O}(1)$ requires at least 6 free parameters. This point will be relevant in forthcoming discussions.

The normalized inelasticity densities $G_{\mathrm{in}}(s,b)$ extracted from the corresponding fits to $\mathrm{d}\sigma_{\mathrm{el}}/\mathrm{d}t$ for ISR and LHC energies are shown in Fig. 2.3. Note that by construction $G_{\mathrm{in}}(s,b) < 1$ as required by the unitarity condition. Analyzing the behavior of $G_{\mathrm{in}}(s,b)$ in Fig. 2.3 a novel, striking feature of p+p interactions emerges: at high energies, the inelasticity density of the collision does not reach a maximum at zero impact parameter. Rather, peripheral collisions, where the effective geometric overlap of the colliding protons is smaller, are more inelastic or, equivalently, are more effective in the production of secondary particles than central ones. This phenomenon, not observed before at lower collision energies, is not an artefact of our specific parametrization for $T_{\mathrm{el}}(s,t)$. Actually, identical conclusions have been achieved employing completely different techniques and it has been referred to as *hollowness* [58] or *grayness* [74, 75, 76] of proton-proton collisions by the authors of the first analyses where it was identified.

This exciting phenomenon is not exempt from a certain degree of skepticism. Firstly, from Fig. 2.3 it could be argued that an effect at the level of 10^{-3} can be defined, perhaps, as subtle. Further, it has been discussed in [78] that the accuracy of the experimental data at LHC is still not enough to completely discard the no-hollowness hypothesis i.e. the deviation is in the scope of experimental error bars. Efforts have been made to refute the previous arguments. On the one hand in [73], it was unveiled via the inverse scattering method that the shallow minimum in the 2-D inelasticity density reflects a much more pronounced hollow in 3-D. This idea can be easily understood as the longitudinal integration effectively covers up the hole. Then, a flattening of G_{in} at low impact parameters is already signaling a robust depletion in the 3 dimensional scenario. Besides, an alternative approach to the hollowness phenomenon was suggested in [79]. In this work the elastic scattering amplitude at low values of b is parametrized as purely imaginary and Gaussian

$$\mathrm{Im}T_{\mathrm{el}}(s,t) \propto Ae^{-b^2/B}, \;\; A = \frac{4\sigma_{\mathrm{el}}}{(1+\rho^2)\sigma_{\mathrm{tot}}}, \;\; B = \frac{(1+\rho^2)\sigma_{\mathrm{tot}}^2}{16\pi\sigma_{\mathrm{el}}}. \tag{2.17}$$

The obtained inelasticity density from Eq. (2.17) develops a maximum at a non-zero value of b whenever $\sigma_{\mathrm{el}} > \sigma_{\mathrm{tot}}/4$, which is the case for the LHC collision energies. Another criticism that has been recently made [80] is related, precisely, to the role of the real contribution to the scattering amplitude. Since the real part of the elastic scattering amplitude is much smaller than the imaginary part it is often neglected in the unitarity condition Eq. (2.13). Then, the authors discuss how this *approximate* definition of the inelasticity density, that doesn't longer have a probabilistic interpretation, may develop a hollow but emphasize that when the real part is restored it enables to avoid the *hollowness paradox*. This argument, although interesting, does not apply in our case as a non-zero real part is included in the fits. All in all, the best way to overcome these critics would be the release of new data at 13 TeV with a more extended range of t, to reduce the uncertainties on the Fourier transformation in the low b region. This analysis would definitely help to shift our view of the current experimental observations from suggestive to certain.

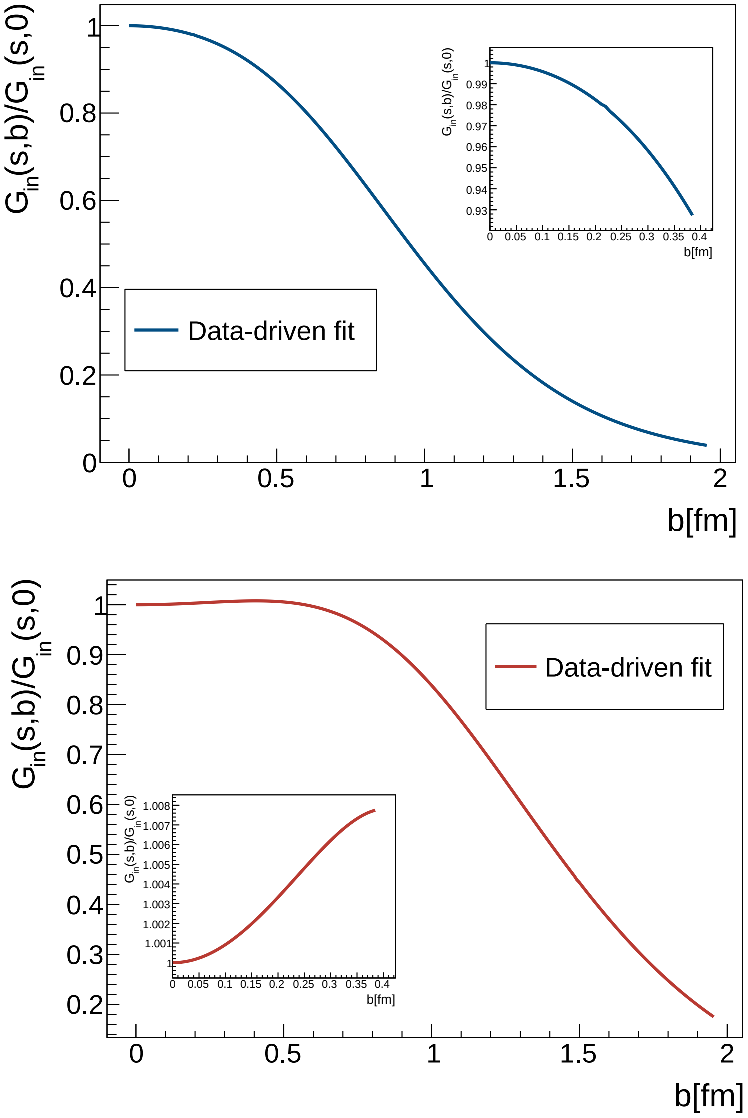

Figure 2.3: Normalized inelasticity density for $\sqrt{s} = 62.5$ GeV (top) and $\sqrt{s} = 7$ TeV (bottom). Sub-pannels: zoom in the region of small b.

Taking everything into consideration, the onset of the hollowness effect entails a change in the derivative of the inelasticity density at $b = 0$, a highly non-trivial dynamical feature. In that regard, its description does not amount to a fine tuning of parameters of a successful parametrization of lower energies data but, rather, requires the introduction of new dynamical mechanisms to account for such critical behaviour. The final point, very important for the subsequent section, is that the *hollowness* effect challenges the standard geometric interpretation of proton-proton collisions. In particular, it precludes models where the scattering amplitude is constructed as a naive folding of the density profiles of the two colliding protons regardless how intricate the internal structure of one individual proton may be. This argument holds even in the tempting case of modeling the proton as empty in the middle. More concretely, and following [81], folding implies

$$G_{\rm in} \propto \int {\rm d}^2 b_1 \int {\rm d}^2 b_2 \rho(\vec{b}_1 + \vec{b}/2)\omega(\vec{b}_1 - \vec{b}_2)\rho(\vec{b}_2 - \vec{b}/2), \tag{2.18}$$

where $\rho(\vec{b})$ describes the transverse distribution of componentes in the proton and $\omega(\vec{b}_1 - \vec{b}_2)$ is usually taken as $\delta(\vec{b}_1 - \vec{b}_2)$ in the literature.

If b is small one can perform a Taylor expansion of Eq. (2.18)

$$\begin{aligned} G_{\rm in} \propto & \int {\rm d}^2 b_1 \int {\rm d}^2 b_2 \rho(\vec{b}_1)\omega(\vec{b}_1 - \vec{b}_2)\rho(\vec{b}_2) - \\ & \frac{1}{2} \int {\rm d}^2 b_1 \int {\rm d}^2 b_2 [\vec{b} \cdot \nabla\rho(\vec{b}_1)]\omega(\vec{b}_1 - \vec{b}_2)[\vec{b} \cdot \nabla\rho(\vec{b}_2)] + \ldots \end{aligned} \tag{2.19}$$

and, therefore, $G_{\rm in} = \alpha - \beta b^2 + ...$, has necessarily a local maximum at $b = 0$, in contrast to the phenomenological observation of the hollowness effect. Eq. (2.18) can also be seen in signal processing language as an autocorrelation function i.e. the convolution of a function with a delayed copy of itself. By construction, the autocorrelation function always reaches its peak at the origin. This fact has far-reaching theoretical consequences: an incoherent superposition of scatterings between partonic constituents cannot reproduce the hollowness effect. Therefore, a quantum nature of the scattering process where interference/coherence effects exist should be the underlying mechanism responsible for the striking onset of the hollowness effect. These observations suggest that the scattering problem may be best formulated in terms of subnucleonic degrees of freedom which internal dynamics and correlations should be non-trivial with increasing collision energy. In the next Section a microscopic realization of the hollowness effect is presented.

2.3 Model description

We propose a phenomenological model allowing for the onset of hollowness of proton-proton interactions. Although QCD-inspired, at the same time, is kept at the highest level of simplicity, the reduced number of parameters being directly related with essential dynamical ingredients. An extensive discussion of the main elements, sketched in Fig. 2.4, can be found below.

2.3.1 Gluonic hot spots

The idea that the gluon content of the proton is concentrated in domains of small radius R_{hs}, as compared to the proton electromagnetic radius that controls the valence quark distribution i.e. $R_{hs} \ll R_p$, is strongly supported by theoretical and phenomenological arguments. In lattice QCD, the gauge-invariant two–point correlator of the gluon

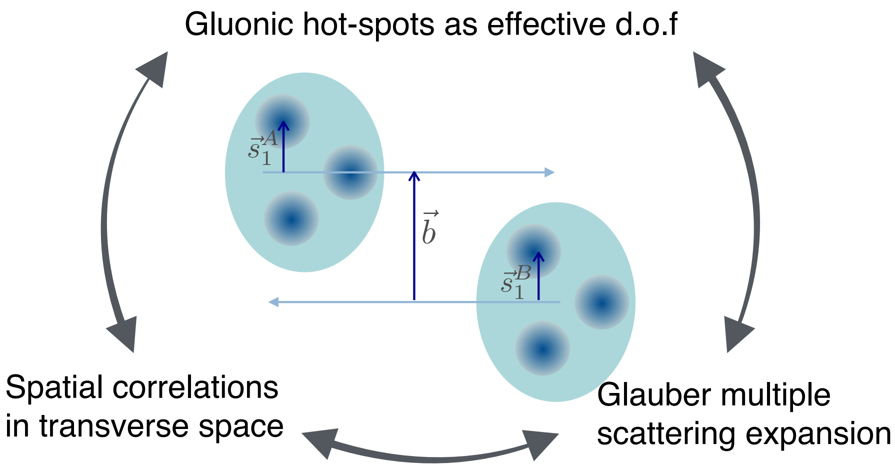

Figure 2.4: Sketch of the main ingredients of our description of proton-proton interactions.

field strength can be computed in the QCD vacuum [82]. It is defined, using the same language of Chapter 1, as

$$\mathcal{D}_{\mu\rho,\nu\sigma}(x) = \langle 0|\mathrm{Tr}\{F_{\mu\nu}(x)\mathcal{G}(x,0)F_{\nu\sigma}(0)\mathcal{G}^\dagger(x,0)\}|0\rangle, \tag{2.20}$$

where x is the spatial coordinate and, the gauge-link is given by

$$\mathcal{G}(x,0) = \mathcal{P}\exp\left(\mathrm{i}\int \mathrm{d}tx^\mu A_\mu(xt)\right). \tag{2.21}$$

An exponential falloff at large distances of Eq. (2.20) i.e. $\mathcal{D}_{\mu\rho,\nu\sigma} \sim \exp(-|x|/\lambda)$ with $\lambda = 0.22$ fm is found. Note that one would expect this distance to be of the order of the hadron size according to the confinement hypothesis. Therefore, an additional scale relevant for the hadron description appears. This result confirms the smallness of the correlation length of the gluon field strengths inside hadrons and suggests a picture of the proton as a collection of domains of high gluonic density. Such domains have been dubbed *gluonic drops* or *hot spots* in the literature.

While the existence of hot spots inside hadrons is widely accepted, the debate on their ultimate dynamical origin remains open. It is commonly assumed that the gluon content of the proton is radiatively generated from valence quarks in DGLAP or BFKL-like cascades (see Eq. (1.26)). In this view, hot spots relate directly to the Fock space of valence partons, for which they would provide an effective description. Thus, a hot spot accounts for a large-x valence quark surrounded by a small-x gluon cloud. However, the question arises of how and why the resulting glue is confined to a region of small radius. In particular, any perturbative emission kernel possesses long range Coulomb-like power tails that allow for successive massless gluon radiation at arbitrary distance scales (as large as the distance from Earth to Cassiopeia) and make the hadron grow exponentially. This mechanism is also known as infrared diffusion. Note that the use of *any kernel* emphasizes the no distinction between linear (DGLAP) or non-linear (BK) kernels. Indeed, in the derivation of the BK equation (see Eq. (1.49)) translational invariance, $N(r,b) \to N(r)$, or, in other words, an homogenous and infinitely large target is assumed. Then, the scattering probability solution of this equation saturates, $N(r) \leq 1$, but locally in impact parameter space. However, once the b-dependence is considered the full BK evolution generate power like tails in $N(r,b)$ that enable scatterings to happen at arbitrarily large impact parameters. This can be directly deduced from the shape of the unintegrated gluon distributions given by Eq. (1.61). In short, perturbative saturation effects cut-off the fast growth of partonic densities but only suppress and not stop the transverse expansion. Phenomenological approaches tackle this problem by including non-perturbative gluon masses that modify the Coulomb behavior to a screened Yukawa one in order to mimic confinement effects. In turn, the intrinsic non-perturbative nature of glue drops has been advocated in [83, 84, 85, 86].

An interesting description of gluonic hot spots within a weak coupling approach motivated by the power-like growth of the total p+p cross section with energy was presented in [87]. The main idea is the following. Two mechanisms are responsible for the rise of $\sigma_{\rm tot}$ when the energy of the collision increases in a perturbative framework: the fast growth of partonic densities and the transverse diffusion. The former is counteracted by saturation effects that kick in when the density reaches a critical value as discussed in Sec. 1.3. The transverse swelling overcomes the non-linear finite density corrections and becomes the leading method driving the evolution of $\sigma_{\rm tot}$ in dense systems. Further, the expansion in the transverse plane is power-like due to the perturbative Coulomb like gluon fields. Thus, it successfully conforms with the experimental observation on $\sigma_{\rm tot}(\sqrt{s})$. To avoid the natural cut-off imposed by confinement, hadrons themselves should be conformed by dense gluonic regions with a radius smaller than the confining scale of QCD. The chosen scale for these glue drops is the radius of the constituent quarks that can be estimated to be ~ 0.3 fm according to diffractive J/Ψ production [88] and double dijet production [89]. Still, non perturbative effects are needed in the asymptotia i.e. when the size of the hot spots reaches the hadronic scale, in order to cease the power-like growth and respect the Froissart bound. Long-range phenomena may be also responsible for the building up of these gluonic clouds around each valence quark at low energies. However, once these glue drops are formed, their energy evolution is strictly driven by perturbative branchings.

In this work we remain agnostic in this debate and simply assume that hot spots are adequate degrees of freedom to discuss inclusive proton-proton scattering at high energies. Our main assumption about their dynamical properties is that collisions between hot spots are fully absorptive over distances smaller than their radius, R_{hs}. Hence, in our effective description, hot spots appear as small black disks of average radius R_{hs}. We also assume implicitly that their ultimate dynamical origin is correlated to the valence partons, since we model the proton as composed of $N_{hs} = 3$ hot spots. The dependence of our conclusions on the number of hot spots will be examined.

2.3.2 Multiple scattering in Glauber theory

Our goal is to construct the elastic scattering amplitude in proton-proton collisions in impact parameter representation. Note that all along the calculation we neglect the longitud-

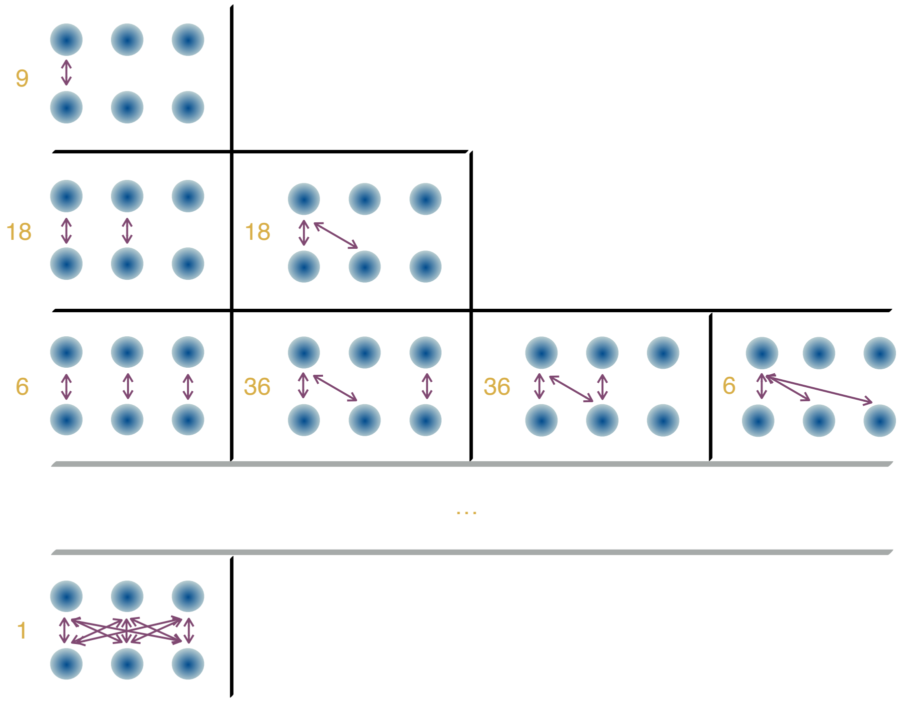

Figure 2.5: Different diagrams contributing to the multiple scattering expansion given by Eq. (2.22). The yellow numbers indicate the combinatorial factor associated to each diagram and the purple arrows the interactions between hot spots. The grey line indicates the 17 omitted diagrams.

inal coordinate due to the Lorentz contraction that suffer the two protons when travelling on the TeV regime. We describe p+p interactions as a collision of two systems, each one composed of three hot spots. According to the Glauber model [90], the natural framework to describe high-energy scattering of composite particles, the elastic amplitude for a collision of hadrons A and B with the hot spots frozen in transverse positions $\{\vec{s}_i\}$ has the form:

$$T_{\rm el}(\vec{b}) = 1 - \prod_{i=1}^{3}\prod_{j=1}^{3}\left[1 - \Theta(\vec{b} + \vec{s}_i^{\,A} - \vec{s}_j^{\,B})\right] , \tag{2.22}$$

where Θ denotes the scattering amplitude of the i-th and j-th hot spots interaction and $\vec{b}$ is the impact parameter of the collisions (see Fig. 2.4). The physical elastic amplitude is obtained after averaging Eq. (2.22) over the transverse positions of the hot spots as given by their probability distributions $D(\vec{s}_1, \vec{s}_2, \vec{s}_3)$ in the projectile and target, A and B:

$$\widetilde{T}_{\rm el}(\vec{b}) = \int \prod_{k,l} {\rm d}^2 s_k^A {\rm d}^2 s_l^B D_A(\{\vec{s}_k^{\,A}\}) D_B(\{\vec{s}_l^{\,B}\}) T_{\rm el}(\vec{b}) \,. \tag{2.23}$$

The main hypotheses underlying Eq. (2.22) are to assume a constant cross section between the constituents and the fact that they do not deviate from their trajectories during the scattering, i.e. the use of the eikonal approximation, both of them justified at high energies. An intuitive way to understand the shape of the multiple scattering expansion given by Eq. (2.22) is to consider the interaction of a projectile A formed by two constituents with a target B with only one. Then, the elastic amplitude computed in a probabilistic way is given by:

$$T_{\mathrm{el}}(\vec{b}) = \Theta_1(1-\Theta_2) + \Theta_2(1-\Theta_1) + \Theta_1\Theta_2, \tag{2.24}$$

where the first two terms correspond to a single scattering while the third one to an interaction in which both constituents interact with the target. Expanding Eq. (2.24) we obtain

$$\widetilde{T}_{\mathrm{el}}(\vec{b}) = \Theta_1(\vec{b}-\vec{s}_1^{\,A}) + \Theta_2(\vec{b}-\vec{s}_2^{\,A}) - \Theta_1(\vec{b}-\vec{s}_1^{\,A})\Theta_2(\vec{b}-\vec{s}_2^{\,A}). \tag{2.25}$$

Therefore, the double scattering term introduces a destructive interference contribution in the elastic amplitude i.e. the scattering as described in the Glauber model contains coherence effects. The different weights in the multiple scattering expansion will play a major role in the description of the hollowness effect. Generalizing Eq. (2.25) to a projectile with A constituents we get

$$\begin{aligned}\widetilde{T}_{\mathrm{el}}(\vec{b}) &= \sum_{i}^{A}\Theta_i(\vec{b}-\vec{s}_i^{\,A}) - \sum_{i>j}^{A}\Theta_i(\vec{b}-\vec{s}_i^{\,A})\Theta_j(\vec{b}-\vec{s}_j^{\,A}) + \sum_{i>j>k}^{A}\Theta_i(\vec{b}-\vec{s}_i^{\,A})\Theta_j(\vec{b}-\vec{s}_j^{\,A})\Theta_k(\vec{b}-\vec{s}_k^{\,A}) \\ &= 1 - \sum_{i}^{A}\left[1-\Theta_i(\vec{b}-\vec{s}_i^{\,A})\right]\end{aligned} \tag{2.26}$$

and, if the target is not elementary but composed by B constituents we recover Eq. (2.22). A diagrammatic representation, pioneered by [91], of some of the different terms contributing to Eq. (2.22) is displayed in Fig. 2.5. Each row corresponds to the k-terms contributing with a weight c^k, indicating the number of equivalent diagrams, to the n-th order of the scattering given by the number of hot spot-hot spot interactions or the number of arrows in the sketch. In our case there are 25 independent diagrams. This diagrammatic scheme is particularly useful to express Eq. (2.23) in a compact manner as

$$\widetilde{T}_{\mathrm{el}}(\vec{b}) = \sum_{k=1}^{25}(-1)^{n+1}c^k T_{\mathrm{el}}^k(\vec{b}), \tag{2.27}$$

where

$$T_{\mathrm{el}}^k(\vec{b}) = \int \mathrm{d}^2 s_1^A \ldots \mathrm{d}^2 s_3^B D_A(\{\vec{s}_i^{\,A}\}) D_B(\{\vec{s}_j^{\,B}\}) \prod_{(i,j)} \Theta_{ij}(\vec{b}+\vec{s}_i^{\,A}-\vec{s}_j^{\,B}) \tag{2.28}$$

involves a 12-dimensional integral. The use of Gaussian functions will allow to analytically solve these integrations, as will be shown in the next Section, thus keeping a tight control over the parameter space.

Further, we have to specify the elastic scattering amplitude between two hot spots separated a transverse distance s_{ij}. As mentioned before, in order to facilitate full analytic calculations, we resort to a Gaussian parametrization:

$$\Theta(s_{ij}) = \exp\left(-s_{ij}^2/2R_{hs}^2\right)(1 - \mathrm{i}\rho_{hs})\,. \tag{2.29}$$

This amplitude can be thought as resulting from the convolution of two Gaussian density distributions for a single hot spot, each of radius R_{hs}. The maximum value of the absorptive part of Eq. (2.29), unity, is reached for head-on collisions. Although hadronic amplitudes are expected to be mostly imaginary at high energies, we allow for a constant real to imaginary ratio ρ_{hs} in order to match the non-zero values measured experimentally. It should be noted that all the energy dependence in our model has been left implicit so far. As discussed below, it is encoded in the transverse swelling of the hot spot radius R_{hs} with increasing collision energy.

To finalize this Section, we would like to mention that considering the number of hot spots as a variable quantity has a dramatic impact on the number of terms spanned by the multiple scattering expansion. For example, the full numerical calculation for $N_{hs}=4$ is an order of magnitude heavier than for $N_{hs}=3$ given the increased dimensionality of the problem. Therefore, motivated by the results concerning the onset of the hollowness effect, presented in the following sections, we have performed the calculation just for $N_{hs}=2,3$.

2.3.3 Spatial correlations

The last element to compute the elastic scattering amplitude in impact parameter space is the spatial distribution of the hot spots. The general structure that we shall consider for the joint probability distribution for the transverse positions of hot spots inside a proton has the following form:

$$D(\{\vec{s}_i\}) = C\left(\prod_{i=1}^{3} d(\vec{s}_i; R)\right) \times f(\vec{s}_1, \vec{s}_2, \vec{s}_3), \tag{2.30}$$

where C is a normalization constant to ensure that the probability distribution is normalized to unity: $\int\{\mathrm{d}^2 s_i\} D(\{s_i\}) = 1$[3]. The next term corresponds to the product of three uncorrelated probability distributions for a single hot spot, $d(\vec{s}_i)$. In order to facilitate a full analytical calculation of the scattering amplitude, we shall assume them to be of a Gaussian form:

$$d(\vec{s}_i; R) = \exp\left(-s_i^2/R^2\right), \tag{2.31}$$

where R is the average radius of the d distribution. It should not be confused with the proton radius itself R_p. Finally, all the correlation structure is encoded in the function f

---

[3] $C = \dfrac{(1+\mu)(1+2\mu)(1+3\mu)^2}{2\mu^3(2+3\mu)\pi^2 R^4}$.

which, by definition, is not factorizable in the spatial coordinates of the hot spots. We write

$$f(\vec{s}_1,\vec{s}_2,\vec{s}_3)=\delta^{(2)}(\vec{s}_1+\vec{s}_2+\vec{s}_3)\prod_{\substack{i<j\\i,j=1}}^{3}\left(1-e^{-\mu|\vec{s}_i-\vec{s}_j|^2/R^2}\right). \tag{2.32}$$

The δ-function in Eq. (2.32) ensures that the hot spots system is described with respect to the proton centre of mass, thus preventing it from acquiring unphysical transverse momentum. Next, we implement repulsive short-range correlations between all pairs of hot spots controlled by an effective repulsive core $r_c^2\equiv R^2/\mu$. In the limit $\mu\to\infty$ we recover the uncorrelated case.

While we have no clear dynamical justification for these correlations, their main role in our calculation is to enforce a larger transverse separation between hot spots with respect to the completely uncorrelated case. Indeed, all realistic models for the electromagnetic nucleon form factors entail non-trivial spatial correlations between the constituent quarks: diquark models, where the proton is envisaged as a bound diquark state interacting with the third quark via gluon flux tubes, depict a rod-like structure of a typical string length $l_s\sim 1.5$ fm. In turn, baryon junction models, where the Wilson lines link the three valence quarks at a junction, yield a more triangular structure of the proton. We argue that three-dimensional realizations of both diquark and baryon junction models, when projected onto the reaction plane, produce a similar correlation structure as the two dimensional repulsive core correlations in Eq. (2.32). This is proved in Fig. 2.6 where we show the average transverse distance between two hot spots yielded by the uncorrelated distribution, corresponding to $\mu\to\infty$, a correlated one with $r_c=0.3$ fm and the corresponding value for two three-dimensional triangular distributions projected onto the reaction plane: equilateral and a highly asymmetric isosceles one, which we take as proxies for baryon junction and diquark models. More quantitatively, average quantities are defined as

$$\langle X\rangle=\int \mathrm{d}\vec{s}_1\mathrm{d}\vec{s}_2\mathrm{d}\vec{s}_3 X D(\vec{s}_1,\vec{s}_2,\vec{s}_3) \tag{2.33}$$

and the probability distribution for the equilateral and isosceles triangles are, respectively:

$$\begin{aligned}D_{\mathrm{BY}}(\vec{r}_1,\vec{r}_2,\vec{r}_3)=C_{\mathrm{BY}}e^{-(r_1^2+r_2^2+r_3^2)/R^2}\delta^{(3)}(\vec{r}_1+\vec{r}_2+\vec{r}_3)\delta^{(3)}(|\vec{r}_1|-|\vec{r}_2|)\\ \times\,\delta^{(3)}(\vec{r}_1\cdot\vec{r}_2+|\vec{r}_1||\vec{r}_2|/2)\end{aligned} \tag{2.34}$$

$$D_{\mathrm{QDQ}}(\vec{r}_1,\vec{r}_2,\vec{r}_3)=C_{\mathrm{QDQ}}e^{-(r_1^2+r_2^2+r_3^2)/R^2}\delta^{(3)}(\vec{r}_1+\vec{r}_2)\delta^{3}(\vec{r}_2-\vec{r}_3) \tag{2.35}$$

where C_{BY} and C_{QDQ} are again normalization constants. Note that in Eqs. (2.34-2.35) the vectors $\vec{r}_i$ are 3-dimensional in opposition to $\vec{s}_i$ in Eqs. (2.30-2.32) which are defined on the transverse plane. As displayed in Fig. 2.6 the results for the probability distribution given by Eqs. (2.30-2.32) provide a good interpolation between the aforementioned, more realistic models of proton substructure.

Albeit the debate on the necessity of spatial correlations inside the proton is very timely it has been previously discussed and analyzed in the case of nuclei [92, 93]. Further, a similar mechanism prevents the ropes in the DIPSY event generator to be in a

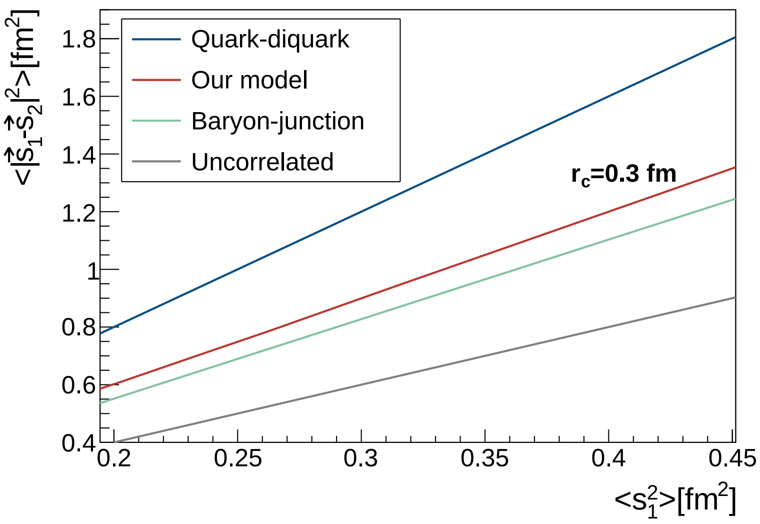

Figure 2.6: Mean transverse separation squared between pairs of hot spots as a function of the mean transverse position for different $D(\{\vec{s}_i\})$ as given by Eqs. (2.30), (2.34) and (2.35).

highly energetic color state [94]. In the Glauber formulation of the scattering process adopted here, the main effect of correlations in the transverse positions is to reweight the contribution of the different terms of the multiple scattering series spanned by Eqs. (2.22-2.23) with respect to the uncorrelated case. For instance, in baryon junction models, terms where the three hot spots in one proton undergo simultaneous scattering with constituents of the target are strongly suppressed, since the three vertices of an equilateral triangle cannot overlap in the transverse plane, unlike the uncorrelated case. The other parameter that controls the amount of effective overlap for different scattering configurations, and hence their relative contribution to the scattering series, is the hot spot radius, that, together with R, r_c and ρ_{hs} are the four parameters of our model.

2.3.4 Multi-dimensional Gaussian integration

All the necessary elements to compute $\widetilde{T}_{\rm el}(\vec{b})$ as given by Eqs. (2.27-2.28) have been described. Now, we shall present a simple algebraic scheme which allows to analytically perform the integration over the transverse hot spots positions exploiting that $D^{A,B}(\vec{s}_1, \vec{s}_2, \vec{s}_3)$ and the elastic scattering amplitude between hot spots, Θ, are Gaussian functions. First, the integrals involved in

$$T^k_{\rm el}(\vec{b}) = \int {\rm d}^2 s^A_1 \ldots {\rm d}^2 s^B_3 D_A(\{\vec{s}^A_i\}) D_B(\{\vec{s}^B_j\}) \prod_{(i,j)} \Theta_{ij}(\vec{b} + \vec{s}^A_i - \vec{s}^B_j) \tag{2.36}$$

can be written in a general way as

$$\mathcal{I}(\vec{b}) = \int \mathrm{d}^2 s_1^A \ldots \mathrm{d}^2 s_3^B \exp(-f(\vec{s}_1^A, \ldots, \vec{s}_3^B; \vec{b})) \tag{2.37}$$

where

$$\begin{aligned} f(\{\vec{s}_i\}; \vec{b}) &= \sum_{i,j=1}^{6} X_{ij} \vec{s}_i \cdot \vec{s}_j + 2\vec{b} \cdot \sum_{i=1}^{6} Y_i \vec{s}_i + W b^2 \\ &= s^{\mathrm{T}} \cdot X \cdot s + 2z \cdot s + W b^2. \end{aligned} \tag{2.38}$$

Note the vectorial character of s and z although, in the last line, the notation is not explicit for readability purposes. In turn, W is a scalar and X a matrix. The goal is to remove the linear term in Eq. (2.38). This can be achieved by transforming the coordinates as follows

$$\beta = \eta + s, \text{ where } \eta = X^{-1} z. \tag{2.39}$$

Then, one obtains

$$\begin{aligned} s^{\mathrm{T}} \cdot X \cdot s + 2z \cdot s &= (\beta - \eta) X (\beta - \eta) + 2z(\beta - \eta) \\ &= (\beta - X^{-1} z) X (\beta - X^{-1} z) + 2z(\beta - X^{-1} z) \\ &= \beta X \beta - \beta X X^{-1} z - X^{-1} z X \beta + X^{-1} z X X^{-1} z + 2z\beta - 2z X^{-1} z \\ &= \beta X \beta - z^{\mathrm{T}} X^{-1} z. \end{aligned} \tag{2.40}$$

Since X is real and symmetric we may make an unitary and orthogonal transformation, $\mathcal{O}$, on the coordinates β_i which will diagonalize it, giving

$$\begin{aligned} \int e^{-\beta^{\mathrm{T}} X \beta} \prod_{i=1}^{6} \mathrm{d}^2 \beta_i &= \int e^{-\alpha^{\mathrm{T}} \mathcal{O}^{-1} X \mathcal{O} \alpha} \prod_{i=1}^{6} \mathrm{d}^2 \alpha_i \\ &= \int e^{[-\lambda_1 \alpha_1^2 - \ldots - \lambda_6 \alpha_6^2]} \prod_{i=1}^{6} \mathrm{d}^2 \alpha_i \\ &= \frac{\pi^6}{\det X} \end{aligned} \tag{2.41}$$

where λ_i are the eigenvalues of X. Thus the Gaussian integral in terms of the matrix X, the vector Y and the scalar W is

$$\mathcal{I}(\vec{b}) = \frac{\pi^6}{\det X} e^{-b^2 (W - Y^T X^{-1} Y)}. \tag{2.42}$$

The explicit forms of X, Y and W change from diagram to diagram of Fig. 2.5. We have systematically computed the 25 different contributions to $\widetilde{T}_{\mathrm{el}}(\vec{b})$ using the symbolic manipulation program Mathematica [95].

2.4 Parameter space: $\{R_{hs}, R_p, r_c, \rho_{hs}\}$

In order to prove that our model actually accounts for the onset of hollowness effect at high energies, we scan the parameter space looking for the presence of a dip of the inelasticity density at zero impact parameter and a monotonically decreasing scattering amplitude, as dictated by data. In our model, essentially a linear combination of Gaussian functions, both $\widetilde{T}_{\rm el}(s,b)$ and $G_{\rm in}(s,b)$ have extrema at $b=0$ by construction (the first derivative vanishes). Therefore, the growing/decreasing behavior of these two functions is governed by the sign of the second derivative at $b=0$. Consequently, we impose the following conditions:

$$\left.\frac{{\rm d}^2\widetilde{T}_{\rm el}(s,b)}{{\rm d}^2 b}\right|_{b=0} < 0, \tag{2.43}$$

$$\left.\frac{{\rm d}^2 G_{\rm in}(s,b)}{{\rm d}^2 b}\right|_{b=0} > 0\,. \tag{2.44}$$

We shall refer to the region of parameter space that fulfills the two above conditions as hollowness region.

A first, important result is that it is not possible to obtain a growing behaviour of $G_{\rm in}(s,b)$ at zero impact parameter in the absence of non-trivial correlations, i.e for $\mu\to\infty$ in Eq. (2.32) or, equivalently for zero correlation distances $r_c=0$, as long as $\rho_{hs}=0$. In turn, for non-zero values of the correlation length $r_c<R$, we find a wide region of the parameter space compatible with the hollowness effect. In Fig. 2.7 we show the hollowness region –represented as a grey area in the plots– in the (R_p, R_{hs})-plane for $r_c=0.3$ (top), 0.4 (middle) and 0.5 fm (bottom), where we have defined $R_p^2 \equiv R^2 + R_{hs}^2$, the effective proton radius resulting from the convolution of the hot spots distribution with their own density distribution. Results in all plots were obtained with $0.05<\rho_{hs}<0.15$. We observe that the hollowness region enlarges with increasing correlation distance r_c.

In order to ensure the compatibility of our results with other global features of experimental data, we now explore the phase space region of our model that is compatible with the measured values of $\sigma_{\rm tot}$ and ρ at LHC and ISR energies as given by Eqs. (2.16). Upon imposing these further phenomenological restrictions we see how the phase space region phenomenologically compatible with ISR data falls outside the hollowness area. In turn, the subspace compatible with LHC data at 7 TeV fully overlaps with it, both results in perfect agreement with empiric observations. These phenomenologically allowed regions are represented in Fig. 2.7 as blue (ISR) and red (LHC) solid areas. It is also shown the subspace of parameter space –represented as a green area in the plots– compatible with the COMPETE predictions $\sigma_{\rm tot} = 111.5 \pm 10$ mb and $\rho = 0.14^{+0.01}_{-0.08}$ for collision energy 13 TeV which, same as for 7 TeV, is fully contained within the hollowness region. We hence predict that the hollowness effect should also be observed for the collision energy of the Run II at the LHC, 13 TeV, provided the COMPETE predictions hold. We have also tested that there is no hollowness region compatible with the phenomenological constraints when the number of hot spots, N_{hs}, is 2. We then infer that the minimum number of hot spots to describe the onset of the hollowness effect, within our model, is $N_{hs}=3$. Possibly, setting $N_{hs}\geq 4$ our conclusions would also hold.

Regarding the relevant geometric scales for fixed R_p and r_c, the hollowness effect kicks in at some finite value of the hot spot radius R_{hs}. For instance, fixing R_p to the value of the measured proton charge radius $R_p \approx 0.88$ fm the hollowness region starts at $R_{hs} \geq 0.34$ and 0.24 fm for $r_c = 0.3$ and 0.5 fm respectively. These results are in perfect agreement with the chosen hot spot radius discussed in Sec. 2.3.1. We hence conclude that the main dynamical process underlying the onset of the hollowness effect is the transverse diffusion or growth of the hot spots with increasing collision energy, which is the main result of this analysis. Further, the measured growth of the total proton-proton cross section can be simultaneously accounted for by the same mechanism.

	R_{hs}[fm]	R_p[fm]	r_c[fm]	ρ_{hs}
ISR	0.21	0.74	0.3	0.14
LHC	0.27	0.94	0.5	0.15

Table 2.3: Parameters of the G_{in} results presented in Fig. 2.8 for ISR and LHC energies.

For completeness, we show in Fig. 2.8 a comparison between the resulting G_{in} of our model for representative values of the parameters of Fig. 2.7 given in Table 2.3 and the inelasticity profiles extracted from data as described above. The small-b part, the essential region where the hollowness effect take place, is correctly described within our framework in all cases. For large impact parameters the tails of the distributions extracted from our model start to deviate from the data-driven profile of G_{in}. The main reason is the relatively simplicity of the model. The possibility of improving the description of the data by introducing a genuinely soft contribution is explored in the next section.

2.5 Model vs. $\mathrm{d}\sigma_{\text{el}}/\mathrm{d}t$ data

After successfully describing the onset of the hollowness effect it would be desirable to obtain a quantitatively precise description of experimental data on $\mathrm{d}\sigma_{\text{el}}/\mathrm{d}t$. Clearly, data on the differential elastic cross section at ISR and LHC energies exhibit a high degree of structure that requieres a similar degree of complexity in any phenomenological model. Indeed, we have tried to perform fits to data on $\mathrm{d}\sigma_{\text{el}}/\mathrm{d}t$ within our model and with its four free parameters $\{R_{hs}, R_p, r_c, \rho_{hs}\}$ constrained to also reproduce the experimental values of σ_{tot} and ρ. Notice that, with these additional constraints, our model has essentially only two degrees of freedom, which turns out to be insufficient to reproduce the highly precise and structured experimental data. Further, the functional form for the hot-spot distribution functions and their scattering amplitude are chosen to be just Gaussian functions in order to facilitate the exact analytical computation of the Glauber series. This approach allows us to establish a general framework that could, and should, be systematically improved in order to provide a good quantitative description of all available data. In the case of LHC, Fig. 2.9 shows the fit to the differential elastic cross section using our original model. A similar quality fit is found in the ISR case. The parameters are displayed in Table 2.4. According to the $\chi^2/\mathrm{d.o.f} \sim \mathcal{O}(10)$, the quality of the fits is manifestly poor. However, the

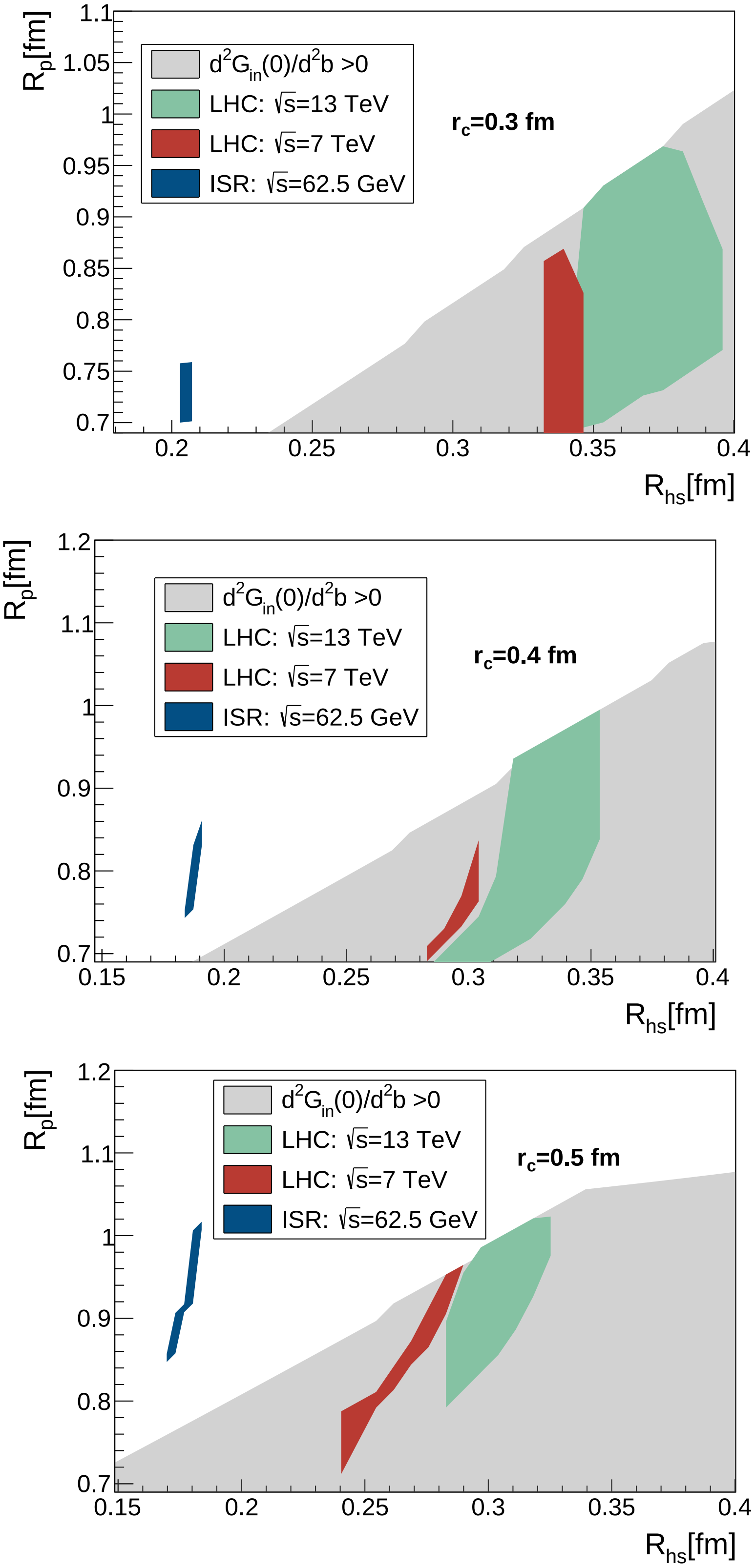

Figure 2.7: Hollowness region (filled-grey) and phenomenologically compatible regions with ISR and LHC data in the (R_p,R_{hs})-plane for $r_c = 0.3$ (top), 0.4 fm (middle) and 0.5 fm (bottom).

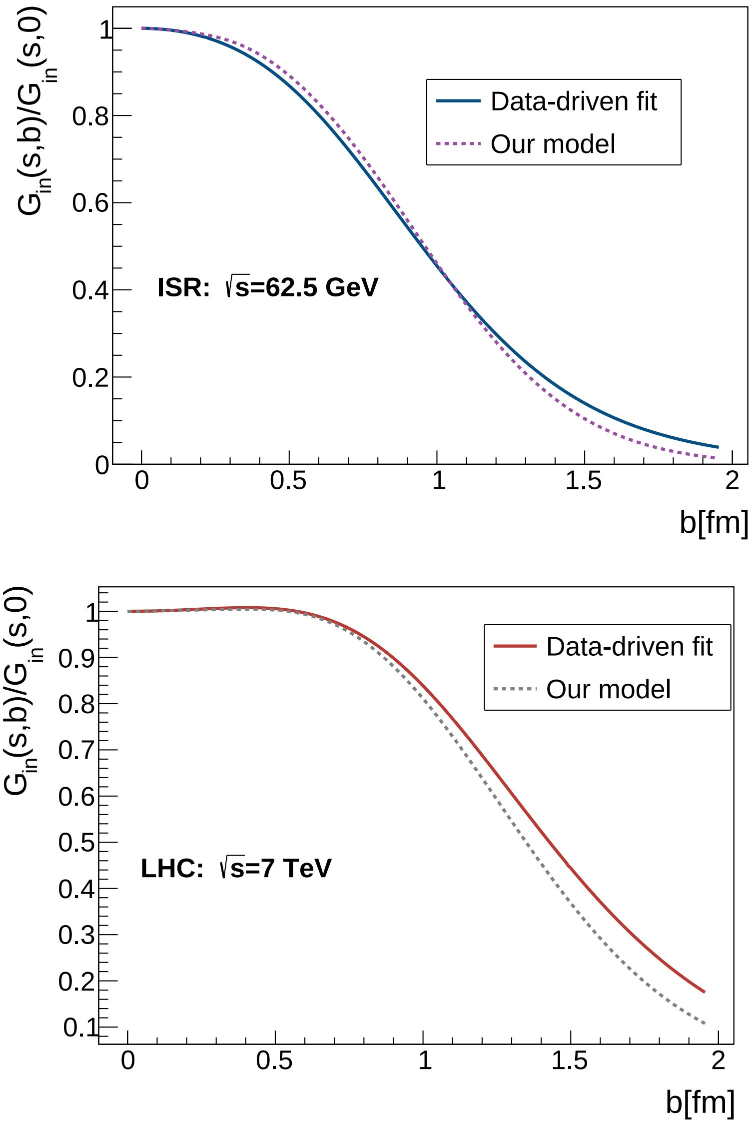

Figure 2.8: Comparison between the values $G_{\rm in}$ extracted from our model (dashed line) and the one obtained after fitting the experimental data (solid line) at ISR (top) and LHC (bottom) energies.

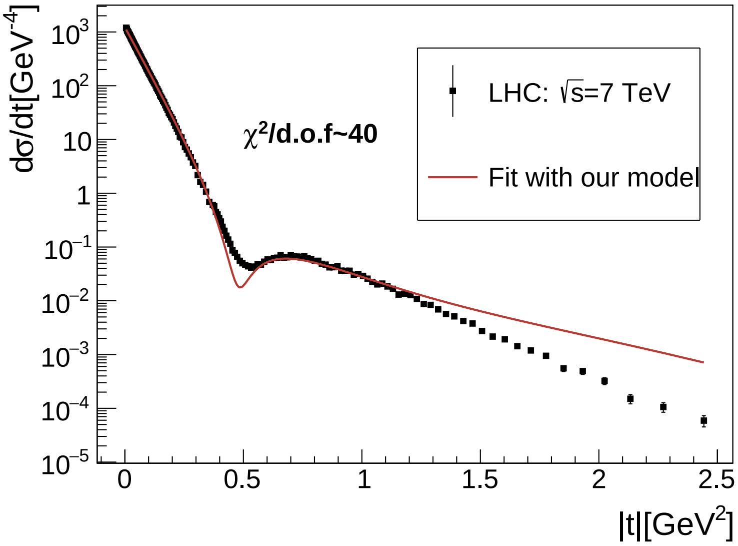

Figure 2.9: Fit to the differential elastic cross section using the parametrization of $T_{\rm el}(t)$ given by our model.

	R_{hs}[fm]	R_p[fm]	r_c[fm]	ρ_{hs}
LHC	0.25	0.99	0.57	0.24

Table 2.4: LHC parameters extracted from the fit with $T_{\rm el}(s,t)$ as computed in our model. The value of ρ is slightly overshot (0.19 vs 0.15).

bulk features of data, namely the slope of the diffraction cone, the position of the minimum and the approximate power-law behaviour at high-$|t|$, are correctly grasped within our very simple model. This result is suggestive that, as we discuss hereafter, a good quantitative description of the data could be attained in case our model was extended and refined.

Arguably, the simplest extension of our model is to allow the scattering amplitude to receive an additional contribution, to which we refer as *soft* as opposed to the supposed *hard* dynamical origin of the hot spot. This new component allows us to adjust the fine detail of the high-$|t|$ behaviour of the scattering amplitude and to obtain a good description of experimental data. Simultaneously, as we argue below, the onset of the hollowness effect is explained in the same terms as in our original model. The scattering amplitude in the extended model now reads:

$$T_{\rm el} = T_{\rm g} + T_{\rm soft}, \tag{2.45}$$

where $T_{\rm g}$ stands for the scattering amplitude resulting from the Glauber series Eqs. (2.27-2.32) and $T_{\rm soft}$ is the new component parametrised in the following way:

$$\mathrm{Im}T_{\rm soft} = A_1 \exp(-B_1 t) \tag{2.46}$$

$$\mathrm{Re}T_{\rm soft} = C_1 \exp(-D_1 t) \tag{2.47}$$

	R_{hs} [fm]	R_p [fm]	r_c [fm]	ρ_{hs}	A_1 [fm^2]	B_1 [fm^2]	C_1 [fm^2]	D_1 [fm^2]
ISR	0.08 ± 0.02	0.82 ± 0.12	0.60 ± 0.05	0.14 ± 0.02	0.87 ± 0.07	0.38 ± 0.09	0.02 ± 0.1	0.10 ± 0.04
LHC	0.26 ± 0.03	0.9 ± 0.08	0.33 ± 0.02	0.17 ± 0.03	1.52 ± 0.8	0.65 ± 0.3	0.06 ± 0.18	0.07 ± 0.02

Table 2.5: ISR and LHC parameters extracted from the fit with $T_{\rm el}{=}T_{\rm g}+T_{\rm soft}$.

or, equivalently, in impact parameter space

$$\mathrm{Im}\widetilde{T}_{\rm soft} = \frac{A_1}{4\pi B_1}\exp(-b^2/4B_1) \tag{2.48}$$

$$\mathrm{Re}\widetilde{T}_{\rm soft} = \frac{C_1}{4\pi D_1}\exp(-b^2/4D_1). \tag{2.49}$$

Adding this new contribution our model have 8 free parameters ($R_{hs}, R_p, r_c, \rho_{hs}$, A_1, B_1, C_1 and D_1) constrained to also reproduce the experimental values of $\sigma_{\rm tot}$ and ρ resulting in 6 degrees of freedom. We have performed global fits to $\mathrm{d}\sigma_{el}/\mathrm{d}t$, finding that a good description of the data is possible, as shown in Fig. 2.10. The fit parameters are presented in Table 2.5. Two important comments are in order:

- First and foremost, the presence of an additional component is needed to attain a good quantitative accurate description of the data at all values of exchanged momentum t but does not change the main conclusions presented in the previous section on the emergence of the hollowness effect. We have checked that the resulting scattering amplitude from all successful fits yields the presence (absence) of hollowness at LHC (ISR) energies. Further, all successful fits yield a non-zero correlation length between hot spots and systematically larger values of the hot spot radius at LHC than at ISR energies. We have not performed a systematic study of the parameter space yielding good fits to data and simultaneously resulting in the onset of hollowness at LHC energies (and its absence at ISR). This would result in a parametric plot analogous to that of Fig. 2.7. Rather, as explained above, we have sampled the parameter space represented in Fig. 2.7 and checked that all the good fits to data obtained within the extended model fall within the right hollowness region of parameter space.

- In the fits, the contribution of the new soft component to the total cross section at ISR and LHC energies is constrained to be less than a 40% and 25% respectively. The relative contribution of the soft term to the total cross section decreases from the ISR to the LHC, thus confirming that the transverse diffusion of the hot spot domains is the main mechanism for the growth of the cross sections with increasing energy.

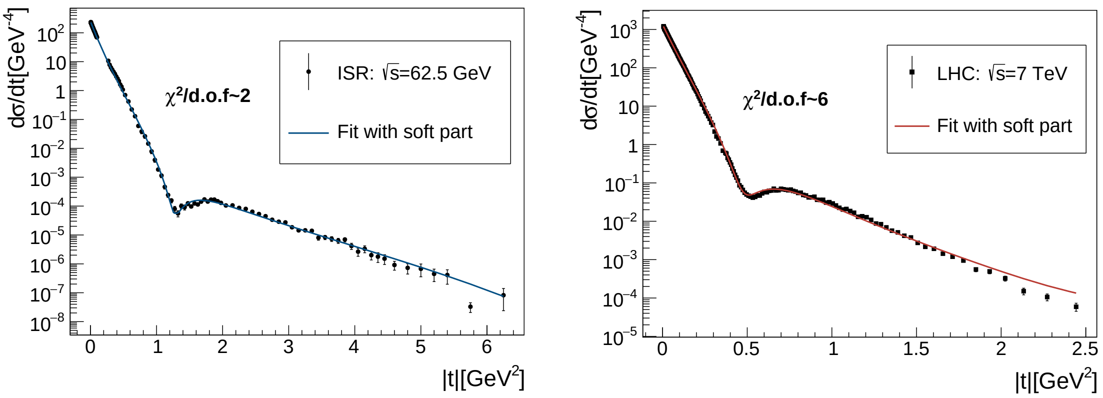

Figure 2.10: Fit to the differential elastic cross section with the $T_g + T_{\rm soft}$ parametrization for the two energies considered in this work. Left: ISR at $\sqrt{s} = 62.5$ GeV. Right: LHC at $\sqrt{s} = 7$ TeV.

The fits to data including a soft component presented above successfully fulfill all the required conditions for the model to be phenomenologically consistent. First, they provide a good quantitative description of the data. Further, the parameters yielded from the best fits to data fall in the correct region of parameter space: no-hollowness for ISR energies and hollowness for the LHC.

2.6 Proton transverse structure as a byproduct of the hollowness effect

Regarding the transverse structure of the proton in terms of gluonic hot spots, describing the onset of the hollowness effect sets stringent constraints on its properties. In a nutshell,

- The number of gluonic hot spots must be at least 3.
- Short-range repulsive correlations between the proton constituents are vital.
- Although we do not provide a dynamical description, the hot spot radius grows with increasing collision energy.

These effects may have observable consequences in other sets of experimental data on proton collisions. Arguably, they could impact significantly the interpretation of data specially sensitive to the initial collision geometry, like the correlation and flow analyses of proton-proton collisions and the possible production of small drops of Quark Gluon Plasma in such collisions, a highly debated topic nowadays. This is the subject of the following chapters.

Everybody talks;
Neon Trees.

3

The quest for the quark-gluon plasma

The natural continuation of the work presented in previous chapters is to explore further consequences of subnucleonic spatial correlations on the properties of the initial state in p+p collisions. As mentioned in the introduction of this thesis, there are several areas of high-energy physics, where the proton structure plays a decisive role. A convenient playground to test the implications of a correlated description of the subnucleonic degrees of freedom of the proton are the initial state properties of high energy proton-proton interactions in the context of quark-gluon plasma (QGP) physics. This new state of matter, to be described below, is expected to be created in extreme conditions when reaching very high temperatures and/or densities. In nature, the early universe and the interior of neutron stars are considered plausible scenarios for the QGP to be formed. Reproducing these conditions in a controlled way is the main goal of high-energy heavy ion collisions (A+A). In this context, the recent analyses of some of the considered golden probes for QGP formation in small collision systems at the LHC (p+p and p+A), thought of as the baseline for heavy ion collisions, have revealed striking similarities to the A+A case. The possibility that QGP droplets are being produced in small collision systems has initiated a lively debate both from a theoretical and an experimental point of view.

In this chapter, a general overview on selected aspects of QGP physics is presented to set the stage. First, the current theoretical paradigm that has been confronted with experimental data on ultra-relativistic heavy ion collisions is described. Then, the signatures of QGP formation are explored in terms of measurable quantities and their interpretation in small collision systems.

3.1 The QCD phase diagram

Up to now we have considered the fundamental degrees of freedom of QCD, quarks and gluons, to be bound inside hadrons. However, another salient feature of QCD, namely asymptotic freedom, tells us that the coupling constant is $\alpha_s(Q^2) \ll 1$ whenever the distance between partons is small or as the exchanged momentum increases. The former situation can be achieved by enlarging the density of the hadronic system i.e. by compressing it, while the latter requires heating the system. Therefore, at sufficiently high densities and/or temperatures the interaction among partons that keeps them confined is weak enough so that they become free [96]. This phase of strongly-interacting matter in which hadron constituents are deconfined and can roam quasi-freely is dubbed the *quark-gluon plasma*.

We have identified the temperature and the density as the relevant variables that make strongly interacting matter show up in distinct phases. In the case of water it is common to represent its different states as a function of the pressure and the temperature. Therefore, analogously to the diagram of water, in Fig. 3.1 we show a sketch ot the QCD phase diagram. As discussed above, the different phases are expressed in terms of the temperature of the system T and a proxy for the density, namely the baryon-chemical potential μ_{B}. The definition of this quantity is the following. The density of baryons, n_B, in the center of the nucleus at $T = 0$ has been measured to be very similar in all nuclei $n_B = 0.16$ fm^{-3}. These baryons inside stable nuclear matter are bound by -8 MeV and therefore, the energy per baryon is $\epsilon/n_B = m_N - 8 \sim 931$ MeV where $m_N \sim 939$ MeV is the rest mass of the nucleon. In general, the basic thermodynamic variables (energy, entropy and number densities and pressure) are related through $\epsilon = Ts + \mu n - p$. Considering a stable nucleus we set the pressure to zero and thus the baryon-chemical potential μ_{B} at $T = 0$ is identical to the energy per baryon in the ground state. Therefore, in the (T, μ_{B})-QCD phase diagram the ground state of nuclear matter sits at $(0, 931)$ MeV. In other words $\mu_B\mathrm{B}$ is a measure of how the energy of the system increases as one adds another baryon to it. At non-zero temperature, if $\mu_{\mathrm{B}} = 0$ the number of baryons and antibaryons is identical while increasing it leads to a baryon dominated system. In the following discussion of Fig. 3.1 we are going to distinguish two different scenarios: the $\mu_{\mathrm{B}} \sim 0$/finite temperature case and a baryon dominated scenario. A detailed description of the QCD phase diagram can be found in [97]. The experimental facilities exploring the different regions of the uncharted (T, μ_{B}) territory will be presented in the next section.

3.1.1 Finite temperature

Nowadays, our universe is baryon dominated. Antibaryons are created in the laboratory or in cosmic ray cascades. This has not always been the case. After the Big Bang and below the electroweak scale ($T < 100$ GeV) the early universe was a relativistic gas of Standard Model particles with the same amount of quarks and antiquarks i.e. $\mu_{\mathrm{B}} = 0$. Given the high temperatures and energy densities, the QCD degrees of freedom formed a quark-gluon plasma. Then, the accelerated expansion made the system cool down and our universe experienced a phase transition from the QGP to hadronic matter in its first microseconds of existence. Note that if the number of baryons and antibaryons would have been exactly the same they would have annihilated each other and only photons would exist. Understanding the mechanism that generated the small excess of matter over antimatter is one of the main open issues of the Standard Model.

It is the goal of modern experiments to reproduce under controlled conditions the high temperatures of the early universe and access the QGP phase. For this purpose, heavy-ion collision at ultra-relativistic energies are performed. In these collisions, the baryo-chemical potential is close to zero. This is so because the nuclei travelling close to the speed of light deposit energy in the reaction centre ($y \sim 0$) and then pass through each other. This deposited energy is converted, due to baryon number conservation, in an equal amount of baryons and antibaryons. As will be discussed below large μ_{B} can be achieved by reducing the beam energy.

From a theoretical point of view, the partition function contains all the information about the thermodynamic properties of the system. In the grand canonical ensemble, using the same notation as in Eq. (1.43), it reads

$$\mathcal{Z}(T, V, \mu) = \int [\mathcal{D}q][\mathcal{D}\overline{q}]\mathcal{D}A e^{i\int \mathrm{d}^4x \mathcal{L}_{\mathrm{QCD}} + \mu\mathcal{N}} \tag{3.1}$$

where μ is the quark chemical potential and $\mathcal{N} \equiv \overline{q}\gamma_0 q$ [97] is the number density operator associated with the quark number conservation. We can talk indistinctly about quark chemical potential or μ_{B} as they are related by a scaling factor given by the number of quarks. By performing elementary algebraic operations with the grand partition function one obtain all thermodynamic quantities such as the pressure p, the energy density ϵ or the equation of state $p(\epsilon, \mu_B)$ (EoS). Then, phase transitions are determined by studying discontinuities of the derivatives of the pressure at a given (T, μ). For example, a first order phase transition, like water turns into vapor, implies discontinuities in

$$\left.\frac{\partial p}{\partial T}\right|_{\mu}, \ \left.\frac{\partial p}{\partial \mu}\right|_{T} \tag{3.2}$$

while the pressure is continuous at the point (T, μ). In general, a n-th order phase transition is described by the discontinuity of $\partial^n p$, while p and its $n-1$-derivatives are continuous. Instead, a *cross over* transition, like butter melts in a pan, occurs when thermodynamic properties change rapidly in a narrow range of (T, μ) while the pressure and its derivatives remain continous. A full analytical solution for Eq. (3.1) would allow to draw firm conclusions on the QCD phase diagram but, once more, non-perturbative effects hamper the task. However, the QCD partition function can be computed in lattice QCD (see Section 1.2.4) at $\mu_{\mathrm{B}} = 0$. At finite baryon densities the fermionic part of the Euclidean action induces an oscillatory behavior in the integral of $\mathcal{Z}$ that invalidates the usual Monte Carlo sampling. To circumvent the so-called *sign problem* is one of the actual challenges of the lattice QCD community [98, 99]. Nevertheless, when $\mu_{\mathrm{B}} = 0$ lattice QCD has demonstrated a smooth cross-over transition from hadronic degrees of freedom to the QGP phase as it is displayed in Fig. 3.2. We see how energy and entropy densities grow very fast in the region around the cross-over temperature that has been determined to be in the range $T_c = 154 \pm 9$ MeV or, equivalently, at a critical energy density of $\epsilon = 0.34 \pm 0.16$ GeV/fm^{-3} [100]. This can also be interpreted as an explosion in the number of degrees of freedom of the system, when moving from a hadron gas to the QGP phase.

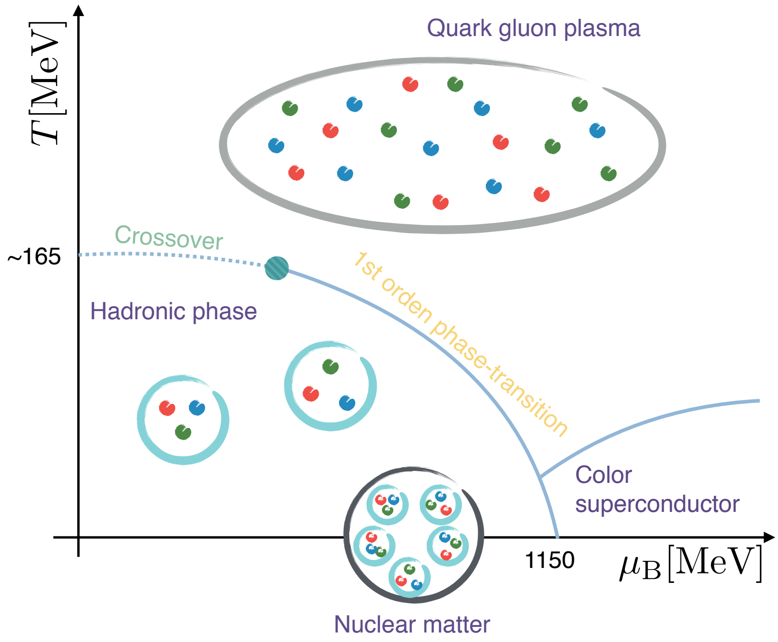

Figure 3.1: Sketch of the QCD phase diagram.

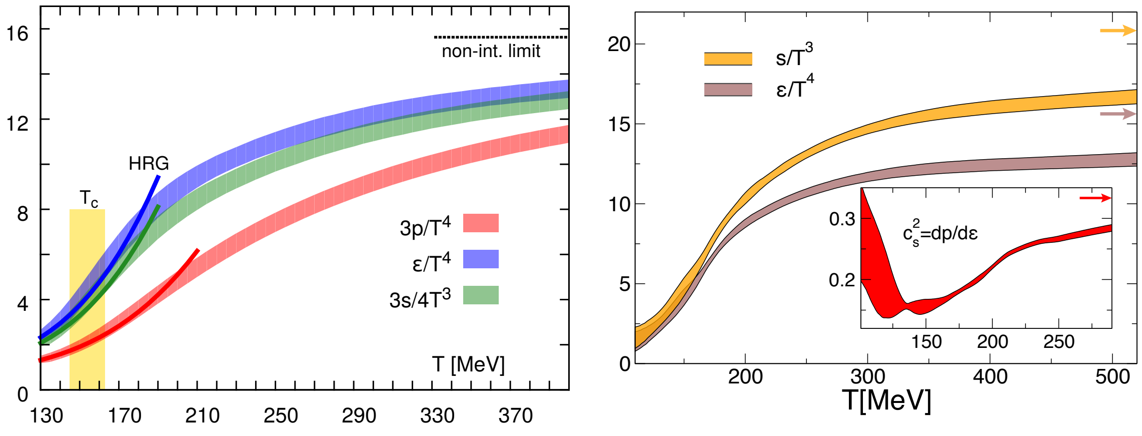

Figure 3.2: Left: pressure, energy density and entropy density as a function of the temperature. The vertical band marks the crossover region. Figure extracted from [101]. Right: energy and entropy densities. The subpanel shows the speed of sound. In both plots the arrows indicate the Stefan-Boltzmann limit i.e ideal massless gas limit. Figure extracted from [102]. This plot shows the agreement between the results obtained by the two major collaborations of the field: HotQCD (left) and Wuppertal-Budapest (right).

3.1.2 Finite density

High energy heavy-ion collisions ($\mathcal{O}(100{-}1000)$ GeV) and lattice QCD are the tools to access the high temperature and low density regime of the phase diagram. Three scenarios have been proposed to explore the high density regime from low to high temperatures.

First, our universe provides a cosmic laboratory where the high density and low temperature phase is expected to play a key role: the interior of neutron stars. The core of these compact stellar objects, formed as a byproduct in supernovae collapse, may reach ten times nuclear matter ground state densities with $T \lesssim 1$ MeV. Then, neutron stars lie on the baryo-chemical potential axis in Fig. 3.1. Therefore, measuring neutron stars properties such as their masses and/or radii may provide insights of cold dense QCD matter [103].

To access the moderate temperature regime ($T \sim 100$ MeV) at finite densities ($2 \lesssim \rho/\rho_0 \lesssim 6$), low beam energies ($\mathcal{O}(1-10)$ GeV) and high interaction rates (luminosity) are required. In this case, the collision occurs at the interaction scale of the nucleons i.e. the partons are not resolved. A phenomenon called *baryon stopping* kicks in at these low energies. It is defined as the shift of an incoming baryon rapidity towards $|y| \sim 0$. The baryons of the colliding nuclei are stopped i.e. they do not leave the interaction region immediately but become part of the medium. Thus, the medium produced in low energy collisions has a larger net baryo-chemical potential due to the contribution of the stopped nucleons.

Finally, it has been argued that similar conditions to the ones in heavy-ion collisions at low beam energies are reached in an alternative scenario. This is the early time of neutron star mergers where both the temperature ($T \approx 100$ MeV) and the density ($\rho \approx 3\rho_0$) are high [104]. Interestingly, it has been explicitly shown how the equation of state of nuclear matter in the high (T, μ_{B})-zone can be constrained by analyzing the post-merger gravitational waves frequency spectra [105]. Consequently, the recent discovery by the LIGO Collaboration of this type of gravitational waves [106] may shed light on the structure of the equation of state of nuclear matter. To conclude, two drastically different experimental setups, heavy ion collisions at low beam energies and neutron star mergers share the necessity of an equation of state. This fact suggests potential sinergies between nuclear theory and cosmology communities.

Theoretically, at low T and high baryo-chemical potential several QCD calculations predict a first-order phase transition [107]. The end of this line would mark the position of a critical point. In the viccinity of a critical point large thermodynamic fluctuations, especially fluctuations of conserved quantities (baryon number, electric charge and strangeness) are expected. These critical fluctuations may show up into multiplicity spectra. Therefore, a suggested experimental way to prove this critical behavior is to measure event-by-event fluctuations and higher-order moments of the multiplicity distribution of baryons, charged particles and strange hadrons [108]. Definite conclusions on the existence of the critical point or the first order phase transition have not been drawn yet.

3.2 Heavy-ion experimental facilities

Awaiting for a robust extension of lattice QCD methods to the finite density regime, hadron colliders provide a unique opportunity to gain insight on the QCD phase diagram. It would be desirable to know the reaction trajectories of each of the experimental facilities in the (T, μ_{B})-plane. However, this cannot be done in a model-independent way. This is so because the equation of state is needed to translate energy densities and particle numbers to (T, μ_{B}) values. Given certain experimental conditions, in terms of the collision energy and the atomic number of projectile and target, one may end up in totally different regions of the phase diagram just by changing the EoS. Therefore, in the vanishing baryon-chemical potential region, where lattice QCD is applicable, the (T, μ_{B})-$(\epsilon, n_{\mathrm{B}})$ correspondence is well established while in the $\mu_{\mathrm{B}} > 0$ regime we rely on educated guesses such as Taylor expansions of the QCD action in powers of μ_{B}/T [98] or effective field theories [109].

This section does not pretend to be an extended historical review of all experimental facilities of the field. We are going to focus on two on-going and two upcoming accelerators given their relevance for this thesis and, in general, for QGP physics. The four of them scan distinct regions of the phase diagram and are, therefore, complementary. Moreover, some of the highlights of their experimental programs will be recasted in Sec. 3.4.

3.2.1 Large Hadron Collider

The physics program of the world's largest and most powerful collider, i.e. the LHC, does not only consist in searching for new particles and performing precision measurements of the Standard Model, but includes studies of the quark-gluon plasma. Besides ALICE, a dedicated heavy-ion experiment, ATLAS, CMS and LHCb have joined the quest for the QGP. The available collision systems and the centre of mass ($\sqrt{s}$) or per nucleon pair ($\sqrt{s_{\mathrm{NN}}}$) energies are summarized in Fig. 3.3. Note that the difference in the maximum energies reached in the proton beam with respect to the nuclear one is due to the charge-to-mass ratio Z/A. In short, neutrons inside the nucleus act like a burden for acceleration purposes. Recently, a Xenon run was performed during 8 hours in order to test the sensitivity of heavy-ion observables to the geometry of the collision. The corresponding results have not been released yet.

The main distinctive feature of the LHC is its capability of performing three different types of collisions namely p+p, p+A and A+A (ordered by increasing degree of complexity), at the same effective energy. This allows to break the very complicated puzzle of heavy-ion collisions into simpler pieces i.e. disentangle effects that have nothing to do with the QGP formation. In this way, the proton run is used to characterize the elementary collisions involved in bigger systems. Analogously, proton-nucleus collisions allow to study the influence of the nuclear wave function, the so-called *cold nuclear matter* effects. Once both collision systems (p+p and p+A) are under control they are used as a benchmark for the heavy ion run in which final-state effects are related to the QGP medium. This neat and logical procedure is based on the assumption that a hot and dense medium is solely formed in the A+A case. This long-established hypothesis has been spoiled by unanticipated observations as will be discussed in Sec. 3.5.

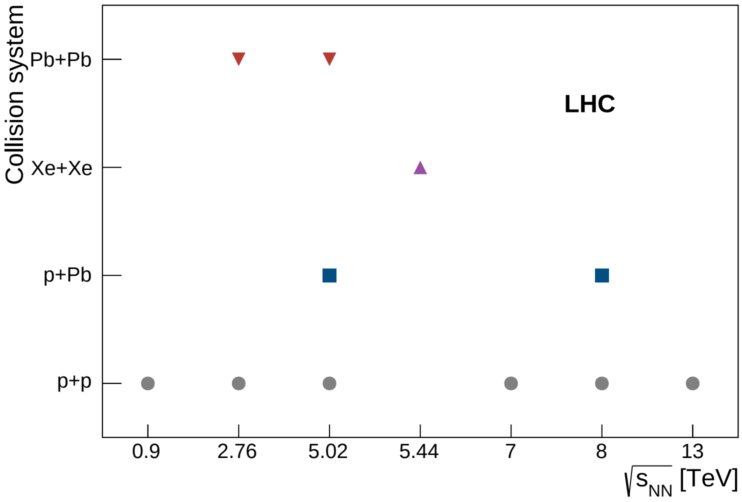

Figure 3.3: Compilation of collisions systems and energies investigated at the LHC. (Pb, Xe) stands for isotopes with atomic numbers (^{208}Pb,^{154}Xe).

3.2.2 Relativistic Heavy Ion Collider

The RHIC in Brookhaven National Laboratory (USA) started running in 2000 becoming the pioneering machine in the experimental study of heavy ions at high energies with Au+Au collisions at $\sqrt{s_{\rm NN}}=200$ GeV. RHIC, in analogy with the LHC, hosts 4 experiments: STAR, PHENIX, PHOBOS and BRAHMS. Nowadays, only the PHENIX and STAR detectors are still active and taking data. The different collision systems and energies available at RHIC are displayed in Fig. 3.4. The most remarkable design characteristic of this facility is the flexibility of the machine that is capable of colliding a wide range of nuclear species at multiple energies. Precisely this versatility lead to the Beam Energy Scan (BES) program whose primary purpose is two-fold. First, pinpoint the location of the critical endpoint and cross the first order phase transition, if it does exist. Second, measure how the QGP signatures turn off, when the energy of the collision decreases below the critical energy density. During 3 years of data taking (2010, 2011 and 2014), deuteron, gold, copper and uranium nuclei were collided at 7 different energies $7.7 < \sqrt{s_{\rm NN}} < 62.4$ GeV (see Fig. 3.4). A second run of the BES, BES-II, is planned to run in 2019-2020 with $\sqrt{s_{\rm NN}} < 20$ GeV in order to collect more statistics given the importance of measuring event-by-event fluctuations of conserved charges for the discovery of the QCD critical point.

RHIC is not only a heavy ion facility but also the most powerful source of polarized proton beams up to date achieving $\sqrt{s}=500$ GeV. This mode is being used to explore the spin structure of the nucleon in the spirit of Chapter 1.

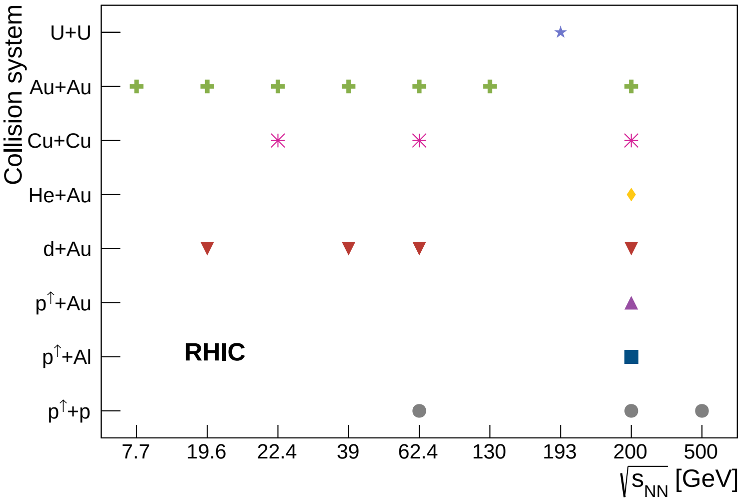

Figure 3.4: Compilation of collisions systems and energies investigated at the RHIC. (He,Al,Cu,Au,U) stands for isotopes with atomic numbers (^{3}He, ^{27}Al, ^{63}Cu, ^{197}Au, ^{238}U). $p^{\uparrow}$ indicates a polarized proton.

3.2.3 Facility for Antiproton and Ion Research

To explore the high-density regime of the QCD phase diagram, the FAIR experiment is currently under construction at the Gesellschaft für Schwerionenforschung (GSI) in Darmstadt, Germany. FAIR is a multi-purpose facility that will cover not only heavy-ion research but also atomic physics, nuclear structure relevant for astrophysics and proton-antiproton collisions. Heavy ion physics at FAIR counts with a fixed target detector, the Compressed Baryonic Matter experiment (CBM) [110]. The goal of the CBM research program is to explore the QCD phase diagram in the region of $800 > \mu_B > 500$ MeV using nucleus-nucleus collisions within an energy interval of $2.7 < \sqrt{s_{\mathrm{NN}}} < 4.2$ GeV. Then, CBM will look for the critical point at even higher densities and smaller collision energies than the Beam Energy Scan program at RHIC. Another particular feature is the interaction rate achieved at CBM that will be at least 2 and up to 5 orders of magnitude larger than in existing and planned low energy heavy ion experiments such as NA61/SHINE at CERN-SPS. This fact enhances the discovery potential of this experiment with respect to others exploring the same region of the phase space.

3.2.4 Electron Ion Collider

The US-EIC will be the world's first e+A collider and will also pioneer the scattering of polarized electrons off polarized protons. There are two laboratories developing plans to host the EIC with the premise of recycling the existing infrastructures: Brookhaven National Laboratory (eRHIC) and Jefferson Lab (MEIC). The machine aims at achieving energies in the range $20 < \sqrt{s} < 175$ GeV for e+p collisions and $40 < \sqrt{s_{\mathrm{NN}}} < 90$ GeV in the e+A case. The beams will consist of diverse species from protons to a Uranium nucleus. Further it can polarize not only electrons and protons but also small nuclei such as d and ^{3}He [21].

The scientific case is based on three main pillars. First, the EIC is a tomographic machine in the e+p mode that offers the opportunity to characterize the structure of the proton in a unprecedented way in terms of TMDs and GPDs as has been presented in Sec 1.2. Turning to the e+A case, the heavy ion community has being facing for a long time the scarceness of data on nuclear parton distribution functions (nPDFs) obtained by the measurements in low-energy fixed-target experiments. This fact translates into large uncertainties that disable a precise study of cold nuclear matter effects needed to properly describe p+A collisions. In [111], it was shown how a significant reduction of these uncertainties could be achieved after including the pseudo data on EIC measurements on the corresponding fits. The third physics prospect of the proposed Electron Ion Collider is to scrutinize the saturation regime by exploring the small-x part of the wave function of nuclei down to $x\sim 10^{-5}$. A region, up to now, truly *terra incognita*.

The European counterpart to the US-EIC, the Large Hadron electron Collider (LHeC), is proposed to be based at CERN [112]. By adding an electron linear accelerator close to the existing LHC tunnel, it would be the only TeV-scale lepton-hadron/nucleus collider ever built reaching a factor of 20 higher centre of mass energy than HERA ($\sqrt{s_{\rm NN}} \approx 0.81-1.3$ TeV). In addition, the electron beam could be polarized. The scientific case of this program would benefit both the particle physics and the heavy ion communities. First, the e+p run will extend the kinematic coverage down to very small longitudinal momentum fraction $x\sim 10^{-7}$ and therefore provide a sound reduction on the uncertainties of PDFs, a crucial element for Beyond the Standard Model searches. In the case of e+A collisions, it would complement the EIC program in the high-energy regime.

3.3 The standard model for heavy ion collisions

Over the last decade, a well-established paradigm on how to describe the 20 fm/c of a heavy ion collision at high energies (RHIC, LHC) has been accepted by the community. The stages of a heavy ion collision are sketched in Fig. 3.5. Contrary to the QCD phase diagram cartoon, the displayed space-time evolution is theoretically motivated. The dynamical description of such a complex process is highly interdisciplinary and it requires the use of diverse theoretical tools: from a CGC picture of the colliding nuclei to the use of kinetic theory in the final stages. The standard model for heavy ion collisions consists of the following components:

- Prior to the collision, the wave function of the colliding nuclei is dominated by small-x gluons. Therefore, they shall be described as sheets of color glass.
- At $t=0$ the nuclei meet each other and the interactions start occuring. Then, at $t=0^+$ the gluons of the colliding nuclei are liberated and form a highly-dense and coherent state out of thermal equilibrium, the so-called *glasma*.

- The large number of interactions between the produced partons lead to a redistribution of energy and momentum that drives the system towards local thermal equilibrium. According to experimental data this process happens in ~ 1 fm/c. How the system thermalizes so rapidly is one of the remaining not fully answered questions of the field.

- This partonic thermal state is nothing but the quark-gluon plasma. In this stage of the evolution, the mean free path is much more smaller than the size of the system and a fluid description is justified. Thus, the energy momentum tensor computed within the CGC serves as an initial condition for the subsequent relativistic hydrodynamic evolution of the fireball. The matching procedure between both formalisms comes, as expected, fraught with subtleties.

- Due to the expansion, the temperature of the system will eventually ($t \sim 10$ fm/c) drop below $T_c = 150 \div 170$ MeV. At this point, bf hadronization kicks in and the equation of state has to be modified from partonic to hadronic degrees of freedom. Then, the system can be described as a hot hadron gas. During this 10 fm/c thermalization dynamics has to outrun the expansion.

- Once in the hadronic phase, the fluid description gradually breaks down. The transition from macroscopic variables (components of the energy momentum-tensor) to a microscopic description in terms of particles (4-momenta) is called *particlization*. The switching criterion is not unique in the literature and varies from a constant temperature of the order of T_c to a critical energy density.

- The lower temperature description of the fireball relies on kinetic theory i.e. the relevant variable is the particle distribution function $f(x,p)$. Hadrons rescatter until the system become so dilute that they do not undergo more interactions i.e. they freeze out at $t \sim 20$ fm/c. After the kinetic freeze out, the hadrons free stream until they reach the detector.

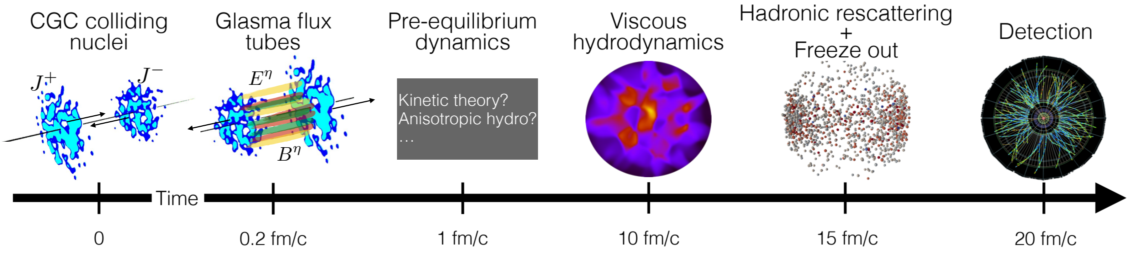

Figure 3.5: The standard model for heavy ion collisions.

More details on each of the stages of high energy heavy ion collisions are presented below.

3.3.1 Initial state and the glasma

Highly energetic nuclei afford an ab-initio description of their structure in terms of the Color Glas Condensate theory (see Sec. 1.3). The large-x degrees of freedom of the two colliding nuclei generate an external current [43]:

$$J^{\mu} = \delta^{\mu +}\delta(x^{-})\rho_{1}(\mathbf{x}_{\perp}) + \delta^{\mu -}\delta(x^{+})\rho_{2}(\mathbf{x}_{\perp}) \tag{3.3}$$

where $\rho_{(1,2)}$ are the color charge densities of projectile and target. This current acts as an external non-dynamical source for the small-x gluon fields via classical Yang-Mills equations $[D_{\nu}, F^{\nu\mu}] = J^{\nu}$. At leading order, observables depend on the solution A^{μ}_{a} that is non-linear to all orders in both ρ_1 and ρ_2 and can be expressed in terms of the associated chromo-electric and chromo-magnetic fields, $(\mathbf{E}_a, \mathbf{B}_a)$. For example, the energy-momentum tensor reads [113]

$$T^{\mu\nu} = \frac{1}{4} g^{\mu\nu} F^{\lambda\sigma} F_{\lambda\sigma} - F^{\mu\lambda} F^{\nu}_{\lambda} \tag{3.4}$$

Prior to the collision ($t < 0$), when Eq. (1.55) is rewritten in the light-cone gauge ($A^{+} = 0$) one obtains the following solution for the gauge field [114]

$$\begin{aligned} &A^{\pm} = 0 \\ &A^{i}_{(A,B)}(\mathbf{x}_{\perp}) = \frac{\mathrm{i}}{g}\mathcal{U}_{(A,B)}(\mathbf{x}_{\perp})\partial_{i}\mathcal{U}^{\dagger}_{(A,B)}(\mathbf{x}_{\perp}) \end{aligned} \tag{3.5}$$

where A refers to the projectile and B to the target. Therefore, the corresponding vectors $(\mathbf{E}_a, \mathbf{B}_a)$ have only transverse components $i = 1, 2$ with respect to the z-axis.

The next step is to compute the gauge field A_{μ} immediately after the two nuclei hit each other i.e. at time $t = 0$. In an Abelian situation, A_{μ} would be simply the sum of the incoming transverse fields of each of the nuclei. However, the intrinsic non-linearities of Yang-Mills equations introduce additional terms. As we have discussed, the two incoming nuclei satisfy, respectively, the conditions $A^{-}(x^{-} = 0) = 0$ and $A^{+}(x^{+} = 0) = 0$. The Fock-Schwinger gauge $A^{\tau} = (x^{+}A^{-} + x^{-}A^{+})/\tau = 0$ (see Appendix B) smoothly interpolates between the light-cone gauge conditions of the incoming nuclei. In this gauge, the solution to the classical Yang-Mills equations of motion at $t = 0$ in terms of the gauge fields of the colliding nuclei reads [115]

$$\begin{aligned} A^{i} &= A^{i}_{A} + A^{i}_{B} \\ A^{\eta} &= \frac{\mathrm{i}g}{2}[A^{i}_{A}, A^{i}_{B}] \end{aligned} \tag{3.6}$$

This solution has a very rich structure. As anticipated, in addition to the intrinsic transverse fields of the colliding color sheets, there is also a longitudinal component of the gauge field A^{η} which extends along the collision axis. At $t = 0$, it can be related to the longitudinal component of the chromoelectric field like $A^{\eta} = -E_{\eta}/2$. This fact suggests a picture in terms of color flux tubes with the endpoints on each of the nuclei and a typical transverse extension radius of $1/Q_s$. Therefore, at the initial time of a high energy nuclear collision the only nonzero components of $F^{\mu\nu}$ are the longitudinal chromoelectric and chromomagnetic fields. Then, the structure of the energy-momentum tensor is $T^{\mu\nu}_{\mathrm{CGC}} = \mathrm{diag}(\epsilon, \epsilon, \epsilon, -\epsilon)$ where the energy density can be computed as [116]

$$\epsilon(t=0) = \frac{1}{2}\mathrm{Tr}(F_{ij}F_{ij}) + 4\mathrm{Tr}(A^{\eta})^2 \tag{3.7}$$

The shape of $T^{\mu\nu}$ in the CGC is comparable to that of a system with an anisotropy in the longitudinal direction i.e $T^{\mu\nu}_{\mathrm{CGC}} = \mathrm{diag}(\epsilon, p_T, p_T, p_L)$ with the transverse and longitudinal pressures given by $(p_T = \epsilon, p_L = -\epsilon)$. At $t=0$, the Glasma is a highly anisotropic state i.e. the ratio p_L/p_T is not close to one. A natural question is, if the time-evolution of the energy-momentum tensor may lead to isotropization and justify the matching with a fluid description.

To compute the evolution of the system at later times, one has to solve the classical equations of motion with Eq. (3.6) as initial conditions. So far, the analytic solution is not known so it has to be computed either numerically or within certain approximations. Note that with increasing time the system expands and becomes dilute in a such a way that the description in terms of gluonic fields breaks down. An estimate of the instant up to which the dynamics can be described in terms of the CGC is $t \sim 1/Q_s \sim 0.2 \div 0.3$ fm/c. This estimate arises from the uncertainty principle: the small-x gluons have predominantly $k_\perp \sim Q_s$ and, consequently, will be released by the collision after $\Delta t \sim 1/Q_s$. Up to $t \sim 1/Q_s$, the numerical solution of classical Yang-Mills (also at NLO [117]), assuming a McLerran-Venugopalan distribution for the color charge densities (see Eq. (1.59), shows that the longitudinal pressure remains at all times much smaller than the transverse one [118]. This leads to an unsatisfactory matching between the CGC description and hydrodynamics at $t \sim 1/Q_s$ [119].

To bypass this fundamental problem several solutions have been proposed. The *bottom-up* thermalization scenario [120] has shown promising results [121, 122]. In general terms, the idea is to introduce an intermediate step between the CGC and hydrodynamics. The natural choice is to describe the system from $t \sim 1/Q_s$ up to t_{hydro} in kinetic theory. Switching from fields to particles at $t \sim 1/Q_s$ is justified as the occupancy numbers are no longer non-perturbative. The input for the effective kinetic theory is the gluon spectrum $dN_g/dyd^2k_\perp$, that is a function of A^μ in the CGC. Then, when the system is sufficiently close to thermal equilibrium, kinetic theory and hydrodynamics provide equivalent descriptions and a smooth matching between them can be performed. Explicit results of two implementations of this scenario are shown in Fig. 3.6. An orthogonal solution to the early thermalization puzzle was proposed in the context of the AdS/CFT correspondance: viscous hydrodynamics may be applicable, even when the pressures are anisotropic i.e. for times as early as $t \sim 0.25$ fm [123, 124]. As a disclaimer, it is important to keep in mind that $\mathcal{N} = 4$ supersymmetric Yang-Mills is a toy model for QCD and henceforth, this result should be considered a valuable insight but not a dynamical solution to the problem of thermalization.

Up to this point we have presented a completely model-independent description of the initial state of high-energy nuclear collisions within the Color Glass Condensate effective theory of QCD. Note that we have omitted so far the energy dependence of the distribution of color charge densities. To extrapolate, in a consistent way, results from

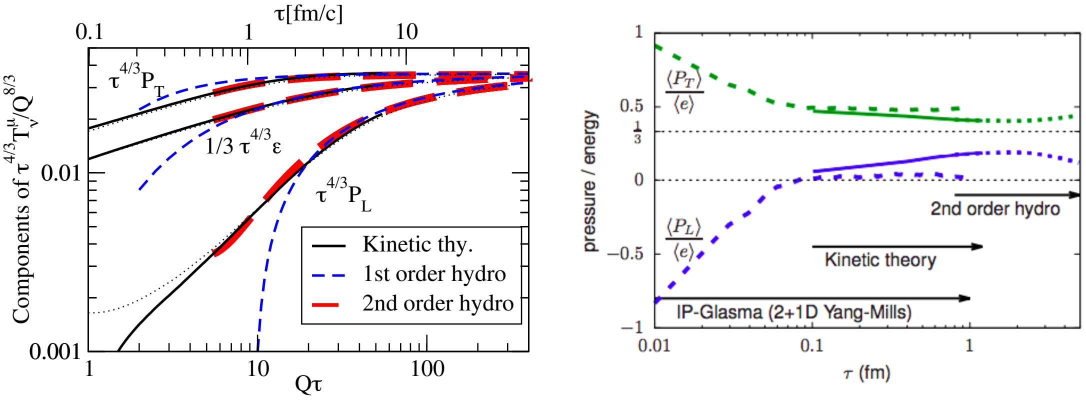

Figure 3.6: Left: Different components of the energy momentum tensor as a function of the proper time. Figure extracted from [121]. Right: Pressure evolution starting from a CGC-inspired model evolved through kinetic theory and matched to hydrodynamics. Figure extracted from [125].

RHIC to LHC energies a full solution of JIMWLK equations is mandatory. However, as we have already mentioned, when including the impact parameter dependence into the evolution equations a regulator/cut-off has to be included ad-hoc to mimic confinement effects. Such an insufficiency of the theory implies some degree of modeling in the nuclear wave functions at small-x to compare with experimental data. The main differences amongst the diverse phenomenological CGC approaches [126, 127] stem from this characterization. To provide a more detailed description on how the initial state modelling is done, we pick IP-Glasma [128], the trendy choice given its very good agreement with a wide range of experimental data, as an example. We stick to a heavy-ion collision where subnucleonic degrees of freedom are not considered. For a discussion of their role see Sec. 3.5. The main ingredients of IP-Glasma can be expressed in an algorithmic fashion as:

- The transverse positions of the nucleons are sampled from a Woods-Saxon distribution

$$\rho(r) = \rho_0 \frac{1 + w(r/R)^2}{1 + \exp((r-R)/a)} \tag{3.8}$$

  where ρ_0 is the nucleon density, R stands for the nuclear radius, a corresponds to the skin depth and w measures the deviation from a spherical shape [129].

- The color charge squared μ_P per unit transverse area of each nucleon is $(x, \mathrm{b}_\perp)$-dependent and proportional to $Q_s^2(x, \mathrm{b}_\perp)$. The saturation scale is obtained from the IP-Sat model. This model describes the dipole-nucleon cross section as [130]

$$\mathcal{N}(x, r, b) = 1 - \exp\left[-\frac{\pi r^2}{2N_c}\alpha_s(\mu^2)xg(x, \mu^2)T(b)\right] \tag{3.9}$$

  where $xg(x, \mu^2)$ is the gluon distribution evaluated at $\mu^2 = \mu_0^2 + 1/r^2$ and $T(b)$ is the impact parameter profile of the proton chosen to be Gaussian. The evolution of $\mathcal{N}$

is done à la DGLAP and the parameters fitted to HERA data. Note that Eq. (3.9) is an extension of the MV-model that incorporates small-x evolution in a completely ad-hoc way.

- The nucleus color charge $\mu_A(x, \mathrm{b}_\perp)$ is obtained by incoherently summing the individual contributions of each nucleon.

- Once $\mu_A(x, \mathrm{b}_\perp)$ is built the color charge density $\rho^a(\mathbf{x}_\perp)$ is sampled adopting the MV-model (see Eq. (1.60)).

- The color charges $\rho^a(\mathbf{x}_\perp)$ are inserted in the classical Yang-Mills equation of motion that are solved numerically using a lattice formulation. Further, an infrared regulator, $\mathcal{O}(m) \sim \Lambda$, that imposes color neutrality at the nucleon level is inserted into the definition of the Wilson line:

$$\mathcal{U}_{(A,B)}(\mathbf{x}_\perp) = \mathcal{P} \exp\left(-ig \int \mathrm{d}x^- \frac{\rho(\mathbf{x}_\perp)}{\nabla_\perp + m^2}\right) \tag{3.10}$$

Given the quantum nature of the colliding objects this procedure is performed on an event-by-event basis. It leads to an initial energy density distribution, Eq. (3.7), displayed on the left side of Fig. 3.7 that is fed into a hydrodynamic evolution. The main feature is the lumpy shape and high degree of granularity of $\epsilon(\mathbf{x}_\perp)$ due to fluctuations on length scales of $\mathcal{O}(Q_s^{-1})$.

The saturation framework, extensively covered in this section, is not the unique accepted description of the initial state in the heavy-ion community. On the right hand side of Fig. 3.7 the initial energy density for a heavy ion collision computed in an alternative paradigm is shown. It relies on a Monte Carlo implementation of the Glauber Model, discussed in Chapter 1. The methodology to extend the analytical framework in a Monte Carlo fashion will not be covered here, but in the following chapter as it constitutes one of the major projects of this thesis.

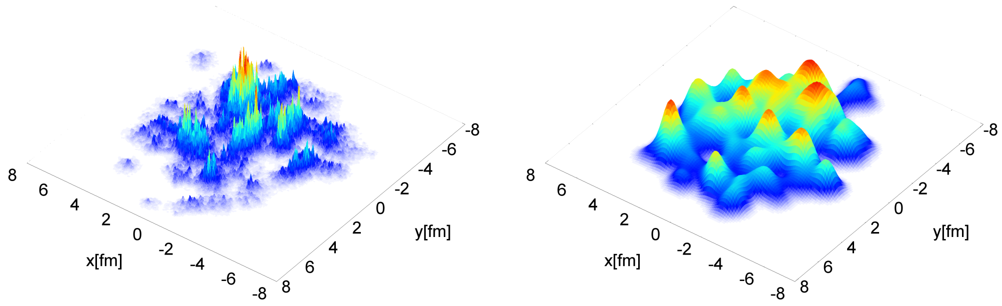

Figure 3.7: Initial energy density in the transverse plane obtained with IP-Glasma (left) and MC-Glauber (right) models. Figure extracted from [128].

3.3.2 Hydrodynamic evolution

Hydrodynamics is an effective theory that describes the dynamics of a fluid in terms of a reduced set of macroscopic variables. The underlying microscopic description is kinetic theory i.e. fluid dynamics follows from kinetic theory in the small mean free path limit. In our case, the fluid under consideration is the expanding and relativistic nuclear fireball. In this section the basic ideas of this framework that have been phenomenologically successful are discussed. Recent developments in the field of hydrodynamical modeling are presented. For a comprehensive review, followed in this thesis, the reader is referred to [131].

A fluid can be characterized locally in space-time by its energy momentum tensor $T^{\mu\nu}(\mathbf{x})$[1] and by the currents j_i^{μ} associated to i conserved charges such as baryon number or electric charge. The equations of motion, in the relativistic case, of the considered fluid are

$$\begin{aligned} &\partial_\mu T^{\mu\nu}(\mathbf{x}) = 0 \\ &\partial_\mu j_i^{\mu}(\mathbf{x}) = 0. \end{aligned} \tag{3.11}$$

These equations demand the conservation of energy, momentum and charge number. $T^{\mu\nu}$ is a symmetric tensor and has, consequently, 10 independent components. Thus, for a system with n charges, there are $10+4n$ independent variables but only $4+n$ are constrained by Eqs. (3.11). Therefore, fluid dynamics is, in its more general formulation, a non-solvable systems of equations. To close the set of equations one has two ways out: reduce the number of unknowns or include $6+3n$ additional equations of motion. The latter is realized in the case of ideal fluid dynamics while the former constitutes an elegant solution when dissipative effects are introduced in the fluid description. Before dealing with both cases it is convenient to decompose $T^{\mu\nu}$ and j^{μ} in their most general form allowed by symmetries:

$$\begin{aligned} T^{\mu\nu} &= \epsilon u^{\mu}u^{\nu} - p\Delta^{\mu\nu} + q^{\mu}u^{\nu} + q^{\nu}u^{\mu} + \pi^{\mu\nu} \\ j_i^{\mu} &= n_i u^{\mu} + \nu^{\mu} \end{aligned} \tag{3.12}$$

where $\Delta^{\mu\nu}$ is a projection operator orthogonal to u^{μ}

$$\Delta^{\mu\nu} \equiv g^{\mu\nu} - u^{\mu}u^{\nu} \tag{3.13}$$

and u^{μ} is an arbitrary normalized time-like vector $u^{\mu}u_{\mu} = 1$. The physical meaning of u^{μ} depends on the reference frame and will be discussed hereafter. As an example, the *Landau frame* is defined as the coordinate system in which $u^{\mu} = (1, \mathbf{0})$. In this reference frame, a physical interpretation of the quantities appearing in Eq. (3.12) is possible e.g ϵ is the energy density, p is the hydrostatic pressure and n_i is the density of particles. Note that with Eq. (3.12) we have just split the $10+4n$ unknowns as follows: n_i depends on n variables, ϵ and p are scalars and $q^{\mu}(3), \nu_i^{\mu}(3n), \pi^{\mu\nu}(5)$.

On general grounds, the microscopic properties of the particles that compose the fluid and its macroscopic description must be related. This is realized in kinetic theory. The

---

[1] Note the use of x as a spatial coordinate.

cornerstone of kinetic theory is the one-particle phase-space distribution function $f(\mathbf{x},\mathbf{p})$ that is nothing else but the Wigner function (see Sec. 1.2). It measures the number of on-shell particles in a given unit of phase space, in this case, of the fluid. Actually, omitting the temporal dependence, $T^{\mu\nu}$ and j^{μ} can be computed as moments of $f(\mathbf{x},\mathbf{p})$ in a local thermal equilibrium state:

$$\begin{aligned} T^{\mu\nu}(\mathbf{x}) &\equiv \int \frac{\mathrm{d}^3p}{E} p^{\mu}p^{\nu} f(\mathbf{x},\mathbf{p}) \\ j_i^{\mu}(\mathbf{x}) &\equiv q_i \int \frac{\mathrm{d}^3p}{E} p^{\mu} f(\mathbf{x},\mathbf{p}) \end{aligned} \tag{3.14}$$

where q_i are charges of the individual particles and E fulfills the relativistic dispersion relation $E = \sqrt{\mathbf{p}^2+m^2}$. Unfortunately, the functional shape of $f(\mathbf{x},\mathbf{p})$ is only known when the system is in local thermodynamical equilibrium, as will be used in the ideal fluid discussion. In a non-equilibrium situation, such as a dissipative scenario, $f(\mathbf{x},\mathbf{p})$ lacks of a dynamical meaning in the context of Eqs. (3.14) and one has to, unavoidably, make approximations. Finally, the one-particle phase-space distribution is not an static object but evolves with time as dictated by the relativistic Boltzmann equation that reads in its covariant form:

$$p^{\mu}\partial_{\mu} f_i(t,\mathbf{x},\mathbf{p}) = \mathcal{C}_i[f_i(t,\mathbf{x},\mathbf{p})] \tag{3.15}$$

where all the interactions (decays and collisions) are encoded in the collision kernel $\mathcal{C}$. The solution of this integro-differential equation in 6+1 dimensions is, in general, not known analytically. The two more relevant cases of fluid dynamics in heavy ion collisions are presented below.

- **Ideal fluid dynamics**: A fluid is said to be ideal, if it is in local thermal equilibrium. The equilibrium phase space distribution reads

  $$f_{\mathrm{eq}}(\mathbf{x},\mathbf{p}) = \frac{g}{(2\pi)^3}\frac{1}{\exp((\mathbf{p}\cdot\mathbf{u}-\mu(\mathbf{x}))/T(\mathbf{x}))\pm 1} \tag{3.16}$$

  where g is a degeneracy factor, $u^{\mu}(x)$ is the average four velocity of the particles, T is the temperature and μ refers to the chemical potential. The $\pm$ relates to the fermionic $(+)$ or bosonic $(-)$ nature of the particle. To compute the macroscopic variables one has to insert f_{eq} into Eqs. (3.14). For an ideal fluid, $T^{\mu\nu}$ and j_i^{μ} are given by

  $$\begin{aligned} T^{\mu\nu} &= \epsilon u^{\mu}u^{\nu} - p\Delta^{\mu\nu} \\ j_i^{\mu} &= n_i u^{\mu} \end{aligned} \tag{3.17}$$

  Comparing Eqs. (3.17) with Eqs. (3.12) we observe how the assumption of local thermal equilibrium significantly reduces the dimensionality of the problem. After setting the dissipative corrections i.e. q^{μ},$\pi^{\mu\nu}$ and ν^{μ} to zero, the system of $4+n$ equations *only* has $5+n$ unknown variables left. The missing element to close the above system is the equation of state $p(\epsilon,n_i)$ that connects the so far general derivation to a particular physical system. Thus, after including the EoS and choosing a set of initial conditions the evolution of the fluid is unambiguously determined by Eqs. (3.11).

- **Viscous fluid dynamics**: The natural extension of the ideal fluid case is to include small departures of thermal equilibrium. In other words, not to set, a priori, the dissipative/viscous terms to zero but force them to be small compared to the equilibrium quantities. This is equivalent to write the distribution function as

  $$f(\mathbf{x},\mathbf{p}) = f_{\text{eq}}(\mathbf{x},\mathbf{p}) + \delta f(\mathbf{x},\mathbf{p}) \tag{3.18}$$

  or, in the hydrodynamic language,

  $$\begin{aligned} T^{\mu\nu} &= T^{\mu\nu}_{\text{eq}} + \delta T^{\mu\nu}, \;\; \delta T^{\mu\nu} \equiv q^{\mu}u^{\nu} + q^{\nu}u^{\mu} + \pi^{\mu\nu} \\ j^{\mu\nu} &= j^{\mu}_{\text{eq}} + \delta j^{\mu} \end{aligned} \tag{3.19}$$

  where $T^{\mu\nu}_{\text{eq}}$ and j^{μ}_{eq} are given by Eqs. (3.17). The question now is how to determine the additional $6+3n$ equations of motion needed to settle the evolution. We are going to outline the two most popular methods in the framework of 2$^{\text{nd}}$ order fluid dynamics.

  The first idea is related to entropy production. Contrary to the ideal case in which the entropy is conserved along the evolution, when dissipative effects are included the entropy increases as dictated by the second law of thermodynamics. This fundamental bedrock is used to derive the evolution equations of the dissipative terms as follows. An ansatz for the entropy four-current is performed [131]:

  $$\begin{aligned} &S^{\mu} = su^{\mu} + \beta q^{\mu} + Q^{\mu} \\ &\quad \text{with } Q^{\mu} = \alpha_0 \Pi q^{\mu} + \alpha_1 \pi^{\mu\nu} q_{\nu} + u^{\mu}(\beta_0 \Pi^2 + \beta_1 \mathbf{q}\cdot\mathbf{q} + \beta_2 \pi^{\nu\lambda}\pi_{\nu\lambda}). \end{aligned} \tag{3.20}$$

  Note that Q^{μ} contains up to second order terms in the dissipative quantities. This fact is essential as when only first order terms are considered the solutions of the evolution equations turn out to be unstable and acausal [132]. Therefore, in order to ensure the principle of non-decreasing entropy one imposes $\partial \cdot S \geq 0$. This leads to a system of equations of motion for the dissipative quantities ($\Pi, q^{\mu}, \pi^{\mu\nu}$) in terms of 8 *transport coefficients* ($\beta_0, \beta_1 \ldots$) [133]. These coefficients, just like the EoS, depend on the microscopic properties of the fluid. They must be obtained from experiment or lattice QCD. The most relevant ones in phenomenological applications are the shear (η) and bulk (ζ) viscosities.

  The second approach relies on kinetic theory. The off-equilibrium correction can be expanded up to second order in momentum like $\delta f(\mathbf{x},\mathbf{p}) = \varepsilon_1(\mathbf{x}) + \boldsymbol{\varepsilon_2}(\mathbf{x}) \cdot \mathbf{p} + \varepsilon_{\mu\nu}(\mathbf{x})p^{\mu}p^{\nu} + \mathcal{O}(p^3)$. Therefore[2], the non-equilibrium one-particle phase-space distribution function depends on 14 independent variables ($\epsilon_1, \epsilon_2, \epsilon^{\mu\nu}$ (symmetric), u^{μ}, T, μ) that can be uniquely mapped to the unknown 14 macroscopic variables ($\epsilon, \Pi, u^{\mu}, q^{\mu}, \pi^{\mu\nu}$) using the kinetic definitions of $T^{\mu\nu}$ and j^{μ} as given by Eqs. (3.14). The equations of motion are then derived from the Boltzmann equation [134]. This is the state-of-the art in hydrodynamic simulations for heavy ion collisions.

In recent years, the necessity to go beyond relativistic viscous hydrodynamics both in the early and final stages of the collision has been clearly established. In these phases,

---

[2] For simplicity, only one conserved charge is considered.

the QGP fireball is highly anisotropic as one would expected given the evident asymmetry between longitudinal and transverse directions in a relativistic heavy ion collision. For the early stages we have already discussed the suggestion of transforming the early non-equilibrated state into a thermalized medium via kinetic evolution of the gluon distribution arising from the CGC. Alternatively, anisotropic viscous hydrodynamics has been proposed to describe far from equilibrium systems [135, 136]. The starting point of this framework is to allow the one-particle distribution function to be momentum anisotropic. In a system that undergoes one-dimensional hydrodynamical expansion $f(\mathbf{x},\mathbf{p})$ reads [137]

$$f(\mathbf{x},\mathbf{p}) = f_{\mathrm{eq}}\left(\frac{\sqrt{p^2+\xi(x)p_z^2}}{\lambda(x)}\right) + \delta f(\mathbf{x},\mathbf{p}) \tag{3.21}$$

where $\xi(x)$ quantifies the momentum anisotropy and $\lambda(x)$ reduces to the temperature in the isotropic limit, $\xi(x) \to 0$. The field is currently under development and the first phenomenological results for bulk observables of heavy ion collisions can be found in [138].

3.3.3 Hadronic rescattering and freeze out

Sooner or later the hydrodynamic description is no longer valid in any of its formulations. Due to the expansion, the system gets colder and dilute leading to a reduction in the number of interactions. Therefore, maintaining local thermodynamical equilibrium becomes unfeasible. Formally, the validity of the fluid-dynamical limit can be quantified by the Knudsen number $\mathrm{Kn} \equiv l_{\mathrm{micro}}/L_{\mathrm{macro}}$. Hydrodynamics applies when $\mathrm{Kn} \ll 1$ i.e. when a clear separation between microscopic length scales (mean free path) and macroscopic scales (related to the gradients) exists. An estimate of the values of Kn reached in heavy ion collisions has been presented in [139].

In practice, the conversion from fluid to particles, also called *particlization*, is performed as follows. First, the transition from partonic to hadronic degrees of freedom is taken care by a change in the equation of state during the hydrodynamic evolution. Note that this does not have to take place at the same instant as the particlization, although is commonly assumed that they occur close in time. Then, single inclusive spectra of particles is obtained from the phase-space distribution function according to the Cooper-Frye formula [140]:

$$E\frac{dN}{d^3p} = \frac{g}{(2\pi^3)}\int_\sigma p^\mu f(\mathbf{x},\mathbf{p})\mathrm{d}\sigma_\mu \tag{3.22}$$

where σ is the freeze-out hyper-surface and $\mathrm{d}\sigma^\mu$ its surface element. Basically, Eq. (3.22) describes the flux of particles through a 3-dimensional hypersurface determined by ensuring a constant temperature, energy density or Knuden number. The number of particle species sampled from the Cooper-Frye prescription is uniquely determined by the degrees of freedom included in the hadronic equation of state.

At this point of the evolution, we are left with hadrons that may collide and decay, although not as frequently as in the early stages, before reaching the detectors. This last step before the free streaming is described within kinetic theory and has non-negligible effects on bulk observables. The combination of hydrodynamics and hadronic rescattering is usually referred to as *hybrid* approaches [141].

The so-called hadronic afterburners describe the non-equilibrium dynamics of hadrons by performing N-body simulations that numerically solve the equations of motion associated to Eq. (3.15) [142, 143]. Hadronic rescattering is the final step in the evolution[3]. Afterwards, particles fly to the detectors. The purpose of next section is to show how to disentangle from this final collection of particles (around 10.000 in a Pb+Pb collision at LHC energies at mid-rapidity) whether a droplet of QGP was created or not.

3.4 QGP signatures

Several observables have been proposed as standard candles for QGP formation. The usual strategy is to compare the outcome of a given observable in a heavy-ion collision with respect to proton-proton considered to be a reliable baseline. To do an apples-to-apples comparison one has to introduce a scaling factor such as the number of nucleon-nucleon collisions in A+A. Then, the differences between both collision systems may be solely attributed to medium effects. Depending on their transverse momentum, p_T, probes are classified into *soft* ($p_\mathrm{T} \lesssim 1 \div 2$ GeV) and *hard* i.e. particles with a perturbatively large p_T. The latter, relatively rare even in the highest energy collisions, do not thermalize but act as tomographic probes. In the following two golden probes of the QGP, one of each sector, will be presented together with their experimental measurement.

3.4.1 Jet quenching

Particles with large transverse momenta are generated in the early stages of a heavy ion collision, well before the QGP is expected to form, via hard scatterings. Being in the large momentum regime, the complete machinery of perturbative QCD is safely and successfully applied. These high energetic parent partons give rise to *jets* via radiation processes i.e. a collimated spray of hadrons in the final state. In the presence of a colored medium the parton propagation is expected to be modified by multiple interactions with the surrounding plasma [144]. The energy loss suffered by a parton during its ride through the medium is referred to as *jet quenching* [145]. Due to this energy loss some of the high energy particles may thermalize becoming part of the medium.

The simplest consequences of jet quenching can be translated into definition of experimental observables. For example, high-p_T hadrons in heavy ion collisions should be suppressed when compared to proton-proton collisions. The observable used to quantify this idea is the *nuclear modification factor*, R_{AA}, defined as

$$R_{AA} = \frac{dN_{AA}/d^2p_\mathrm{T}dy}{\langle N_\mathrm{coll}\rangle dN_{pp}/d^2p_\mathrm{T}dy} \tag{3.23}$$

---

[3] Note that although the theoretical description has been lenghty the actual heavy ion collision takes place in only 20 fm/c.

where N refers to the number of hadrons and $\langle N_{\mathrm{coll}}\rangle$ is the number of binary nucleon collisions. This scaling factor is obtained a priori with Monte Carlo Glauber simulations. In the absence of nuclear modifications this observable would be equal to unity at high p_{T}. At low p_{T}, particle production follows a scaling with the number of participants and not with the number of binary collisions. Therefore, R_{AA} is expected to be lower than one.

In Fig. 3.8 the nuclear modification factor measurements at RHIC and LHC energies are displayed. First, a clear suppression of high-p_{T} hadrons was measured in Au+Au collisions at $\sqrt{s_{\mathrm{NN}}} = 200$ GeV and Pb+Pb at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV. Another interesting feature measured at RHIC is that R_{AA} for photons remains close to unity indicating that electromagnetic probes do not feel the presence of the medium. All in all, the measurement of R_{AA} strongly supports the discovery of the QGP in high energy heavy ion collisions both at RHIC and LHC.

Remarkably, ALICE [146] and CMS [147] have determined the ratio between hadron spectra in p+Pb and p+p. The so-called R_{pA} is intended to disentangle whether medium effects are present in p+A collisions. From Fig. 3.8, we observe the absence of jet quenching i.e. R_{pA} is compatible with one. The straightforward inference is to discard the possibility of creating QGP droplets in small collision systems such as p+A. As will be discussed in Sec. 3.5 this observation is highly challenged by the measurements of a complementary QGP probe: flow harmonic coefficients.

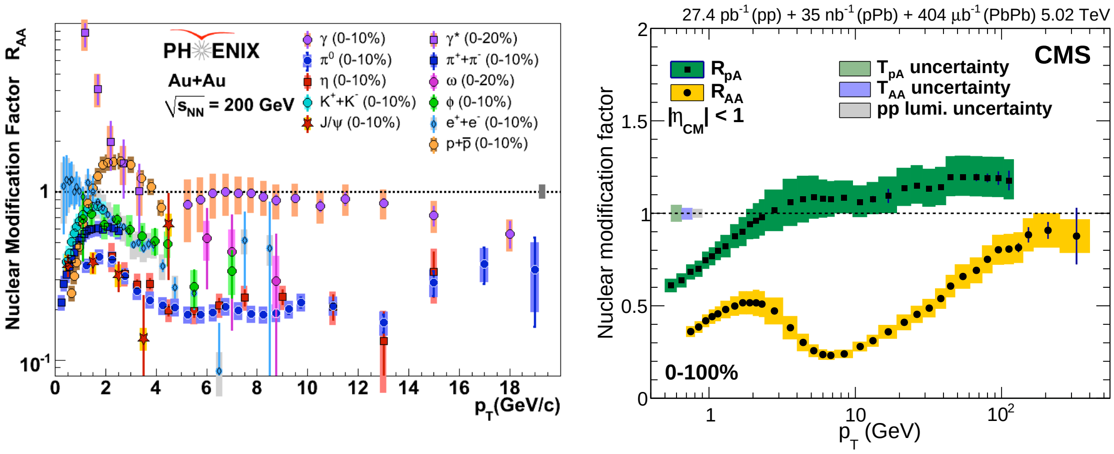

Figure 3.8: Nuclear modification factor at RHIC energies for different particles and centralities (left) and at LHC for Pb+Pb and p+Pb collision systems (right). Figures extracted from [148, 147].

3.4.2 Flow harmonic coefficients

Particle production processes are symmetric with respect to the azimuthal angle, ϕ, whose definition in terms of momentum components of the particle is

$$\phi = \arctan\left(\frac{p_y}{p_x}\right). \tag{3.24}$$

Therefore, one expects an isotropic distribution of particles reaching the detector. However, a key point during our discussion on QGP formation in ultra-relativistic heavy ion collisions is that the system thermalizes. Thus, it can be described in terms of macroscopic variables such as pressure. According to Euler's equation of fluid dynamics i.e.:

$$\frac{\partial \mathbf{v}}{\partial t} = -(\mathbf{v} \cdot \nabla)\mathbf{v} - \frac{1}{\rho}\nabla p \tag{3.25}$$

the pressure, p, leads to a common velocity $\mathbf{v}$ that depends on the density ρ. The common velocity is usually referred to as *collective flow* [149]. This effect has a direct impact on the momentum distribution of particles and constitutes an indirect probe of the QGP transport properties such as its viscosity.

Consider a non-central collision as depicted in the left corner of Fig. 3.9. The overlap region delimits the zone where particles are created, with an isotropic cross section in the azimuthal angle, in the central rapidity region. Note that the shaded area is asymmetric i.e. it is an ellipse with its major axis aligned with the y-direction. This spatial anisotropy leads to a larger pressure gradient in the x-direction. Then, due to the pressure boost, it is easier for the particles to escape in the x-direction than in the y-direction. Consequently, the transverse momenta are preferentially oriented along the same line. Thus, the initial spatial anisotropy of the nuclear overlap region is converted into a final state momentum space anisotropy via large collective pressure gradients during the evolution of the system. This phenomenon translates into a non-flat measured azimuthal distribution of particles, $dN/d\phi$. The whole process is sketched in Fig. 3.9. The plane spanned by the impact parameter and the beam axis, (x, z) in our case, is called *reaction plane*. Notice that the impact parameter fluctuates from one event to the next and so it does the reaction plane.

The azimuthal distribution can be written in the form of a Fourier expansion:

$$\frac{dN}{d\phi} = \frac{v_0}{2\pi} + \frac{1}{\pi}\sum_{n=1}^{\infty} \left[x_n \cos[n(\phi - \psi_{\mathrm{RP}})] + y_n \sin[n(\phi - \psi_{\mathrm{RP}})]\right] \tag{3.26}$$

where ϕ is the azimuthal angle of the particle, ψ_{RP} is the azimuthal angle of the reaction plane and the coefficients are given by

$$\begin{aligned} v_0 &= \int_0^{2\pi} \frac{dN}{d\phi} \mathrm{d}\phi \\ x_n &= \int_0^{2\pi} \frac{dN}{d\phi} \cos[n(\phi - \psi_{\mathrm{RP}})]\mathrm{d}\phi \\ y_n &= \int_0^{2\pi} \frac{dN}{d\phi} \sin[n(\phi - \psi_{\mathrm{RP}})]\mathrm{d}\phi. \end{aligned} \tag{3.27}$$

First, if the azimuthal distribution is normalized $v_0 = 1$. Second, in a symmetric collision the probability of emitting a particle with an angle ϕ is identical to produce it at $-\phi$. Therefore, $dN/d\phi$ is symmetric and all the y_n coefficients vanish. Then, the azimuthal distribution reads

$$\frac{dN}{d\phi} = \frac{1}{2\pi}\left[1 + 2\sum_{n=1}^{\infty} v_n \cos[n(\phi - \psi_{\mathrm{RP}})]\right]. \tag{3.28}$$

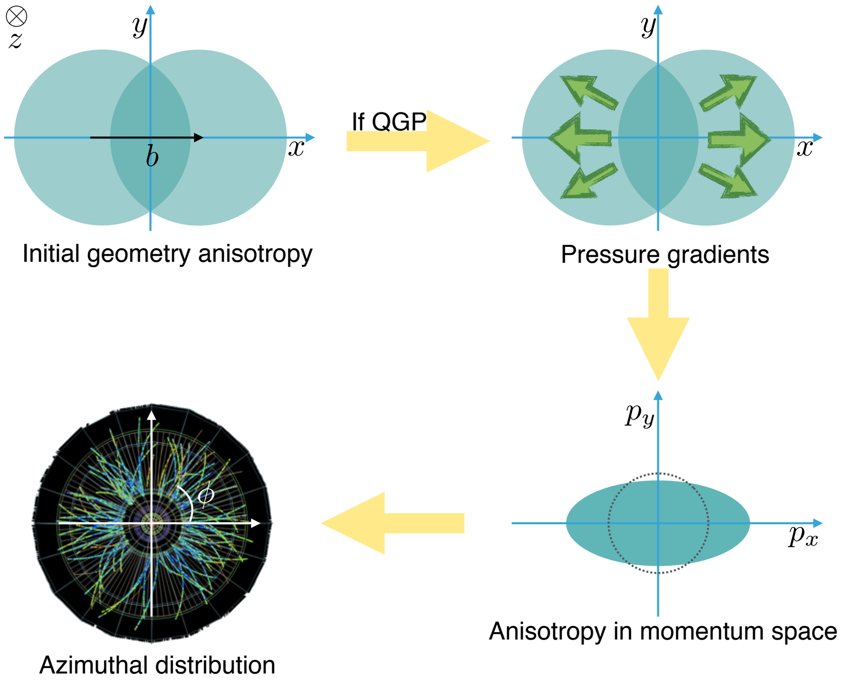

Figure 3.9: Collective flow in a peripheral heavy-ion collision. The impact parameter is directed along the x-direction and z denotes the beam axis.

A more differential definition includes rapidity and transverse momentum dependence, i.e.:

$$E\frac{dN}{d^3p} = \frac{d^2N}{p_\mathrm{T}d^2p_\mathrm{T}dy}\left[1+2\sum_{n=1}^{\infty} v_n(p_\mathrm{T},y)\cos[n(\phi-\psi_\mathrm{RP})]\right]. \tag{3.29}$$

with the so-called *flow harmonic coefficients*, v_n, given by

$$v_n(p_\mathrm{T},y) = \langle\cos[n(\phi-\psi_\mathrm{RP})]\rangle\rangle \tag{3.30}$$

where $\langle\ldots\rangle$ denotes an average over all particles, summed over all events, in the (p_T,y)-bin under study. v_1 is called *directed* flow and represents an overall shift of the distribution in the reaction plane. v_2 and v_3 are referred to as *elliptic* and *triangular* flow respectively, etc.

Although its apparent simplicity, the flow harmonics definition given by Eq. (3.30) entails several complications from an experimental point of view. First, it contains an element that cannot be accessed experimentally: the reaction plane. The easiest way to experimentally estimate the real reaction plane is called the *event-plane* method [150] where the event-plane angle is computed as

$$\Psi_n = \frac{1}{n}\arctan\left(\frac{\displaystyle\sum_{j=1}^{N} w_j\cos(n\phi_j)}{\displaystyle\sum_{j=1}^{N} w_j\sin(n\phi_j)}\right) \tag{3.31}$$

being w_j weights chosen to optimize the reconstruction of the reaction plane. Then, the values of the flow coefficients are estimated as $v_n\{\mathrm{EP}\} \equiv \langle\cos[n(\phi-\Psi_n)]\rangle/\mathcal{R}$. $\mathcal{R}$ is a resolution factor that would be equal to unity if the event plane coincides with the reaction plane. It is commonly calculated by dividing the set of particles into subgroups according to a certain criterion i.e. their rapidity, randomly, etc. If the whole set is divided into two sub-events, A and B, the flow coefficients are estimated as [150]

$$v_n(p_\mathrm{T},y) = \frac{\langle\cos[n(\phi-\Psi_n^A)]\rangle}{\sqrt{\langle\cos[\Psi_n^A-\Psi_n^B]\rangle_E}} \tag{3.32}$$

where the denominator is averaged over the number of events. Certainly, a systematic study of the different choices in this method (weights, number of sub-events,etc.) has to be performed.

The results for the flow harmonic coefficients using the event-plane method both at RHIC and LHC energies are displayed in Fig. 3.10. The experimental data exhibit robust collective flows at both energies. Moreover, as mentioned above, one of the free parameters that can be constrained with flow analyses is the viscosity over entropy ratio, η/s, that characterizes how *perfect* a fluid is, that is, the viscosity is zero for an ideal fluid. The excellent agreement between theory and data displayed in Fig. 3.10 is compatible with values of $\eta/s \sim 0.1$ at RHIC and 0.2 at LHC. These values are an order of magnitude smaller than any other fluid in nature [151]. Then, the QGP is said to be *the most ideal fluid ever observed.* An important comment is in order. Measuring a non-zero value of the flow coefficients do not unambiguously imply a fluid description i.e. the initial state dynamics may be responsible for this observation. In other words, the momentum correlations observed in the final state were there from the very beginning and do not develop during the evolution. Nevertheless, the measured flow strength in ultra-relativistic heavy ion collisions has only be reproduced, up to date, by evolving these initial state momentum correlations with hydrodynamics.

To sum up, jet quenching and collective flow in the expanding fireball, among others measurements, have provided strong indications of the formation of drops of quark-gluon plasma at RHIC and later confirmed at even higher energies at LHC. In the next section the analysis of particle correlations and their theoretical interpretation in smaller collision systems is reported.

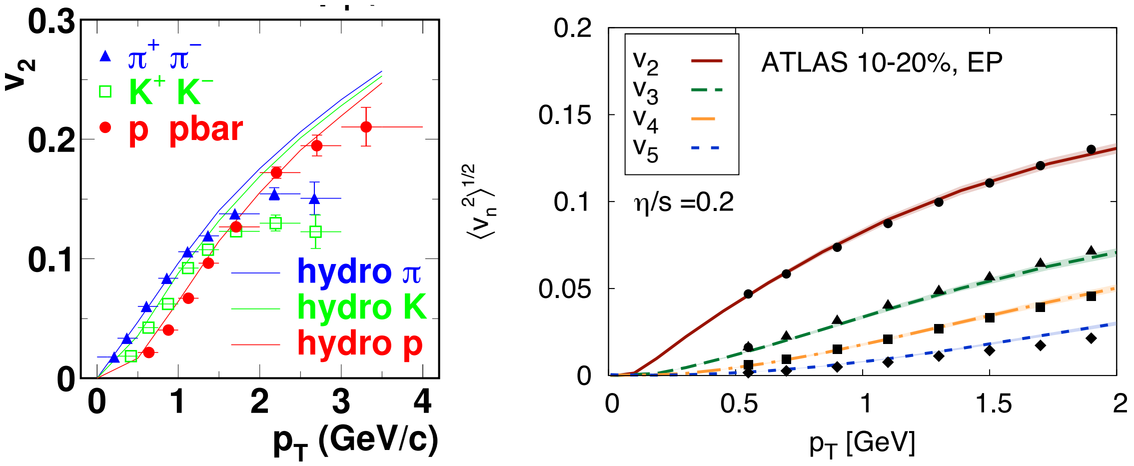

Figure 3.10: Left: elliptic flow measured by PHENIX in Au+Au collision at $\sqrt{s_{\rm NN}}=200$ GeV compared to a hydrodynamic calculation [152, 153]. Right: anisotropic flow coefficients measured by ATLAS in Pb+Pb collisions at $\sqrt{s_{\rm NN}} = 2.76$ TeV compared to IP-Glasma+viscous fluid dynamics simulation [154].

3.5 Hints of collectivity in small systems

3.5.1 Experimental observations

The first discovery at the LHC was the observation in 2010, even before the Higgs boson, by CMS of *ridge* i.e. two-particle correlations at small relative azimuth, $\Delta\phi\sim 0$, which extends over at least several units of $\Delta\eta$ in high multiplicity p+p collisions [155]. This measurement represented a major breakthrough as these correlations could not arise from particles produced in a binary collision i.e due to transverse momentum conservation those would appear in opposite azimuthal hemispheres. These azimuthal correlations were later quantified in terms of flow harmonic coefficients as discussed below.

A fundamental difference between A+A and p+p (or p+A) collisions is the multiplicity of the event. While on a Pb+Pb interaction at the LHC around 10.000 particles are produced at mid-rapidity, the average number of charged particles is ~ 40 in a p+p collision at $\sqrt{s}=7$ TeV. This observation has a direct impact in experimental flow analyses due to the so called *non-flow* contributions, significantly relevant when the multiplicity, N, is low. The event-plane method presented in the previous section relies on a very strong assumption: all momentum correlations in the system are due to anisotropic flow. However, this is not always the case as particles coming from the same resonance decay or contained within the same jet will also have correlated momenta. These sources of correlations, that scale approximately as $1/N$ [156] clearly endanger the flow reconstruction. From a theoretical point of view it is not clear how to compute them. Therefore, the task is to systematically reduce the troublesome non-flow contributions and isolate the collective flow signal. For that purpose, methods using multi-particle azimuthal correlations, where the response to flow and non-flow is better understood, have been developed [157, 158, 159].

To illustrate how these methods works consider the simplest case on which the event-plane method basically relies: two-particle azimuthal correlations i.e. $\langle\cos(n[\phi_1-\phi_2])\rangle$ where ϕ_i is measured with respect to a fixed direction in the detector. The average is performed in two steps: first, over all possible combinations of pairs in one event i.e.

$$\langle\cos(n[\phi_1-\phi_2])\rangle = \frac{1}{N(N-1)}\sum_{\substack{i\neq j\\ i,j=1}}^{N}\cos[n(\phi_i-\phi_j)] \tag{3.33}$$

where autocorrelations are avoided by imposing $i\neq j$. Then, one averages over all events. For a given n harmonic, the measured two particle correlations can be written as:

$$c_n\{2\} \equiv \langle\cos(n[\phi_1-\phi_2])\rangle = \langle e^{(in(\phi_1-\phi_2))}\rangle = \langle v_n^2+\delta_2\rangle \tag{3.34}$$

where δ_2 represents the non-flow contributions and the nomenclature $c_n\{m\}$ has been introduced that refers to the so-called cumulants, which are genuine multi-particle correlations. The standard assumption to neglect δ_2 is only justified when $v_n \gg 1/N^{1/2}$. Further, we have assumed that δ_2 and v_n are uncorrelated. For two-particle cumulants, $c_n\{2\}$, the non-flow correlations can be reduced by requiring a large separation in η between the pairs of particles that enter in Eq. (3.34). In other words, if two particles originate from the same resonance decay they will have a similar value of η. The effect of this rapidity-gap is illustrated on the left side of Fig. 3.11. As expected, non-flow contributions are especially harmful in the low multiplicity regime.

Turning to four particle correlations one finds [163]:

$$\begin{aligned}\langle\cos(n[\phi_1+\phi_2-\phi_3-\phi_4])\rangle &= \langle e^{(in(\phi_1+\phi_2-\phi_3-\phi_4))}\rangle\\ &= \langle v_n^4+(2\times 2)v_n^2\delta_2+2\delta_2^2+\delta_4\rangle\end{aligned} \tag{3.35}$$

where v_n^4 arises from purely flow correlations between the four particles. Its corresponding non-flow term is δ_4 that scales as $1/N^{3/4}$. Additional terms are needed to account for the fact that the four particle correlations may emerge from pairs of two particles correlations. Therefore, non-flow two-particle correlations are included via $2\delta_2^2$ where the factor 2

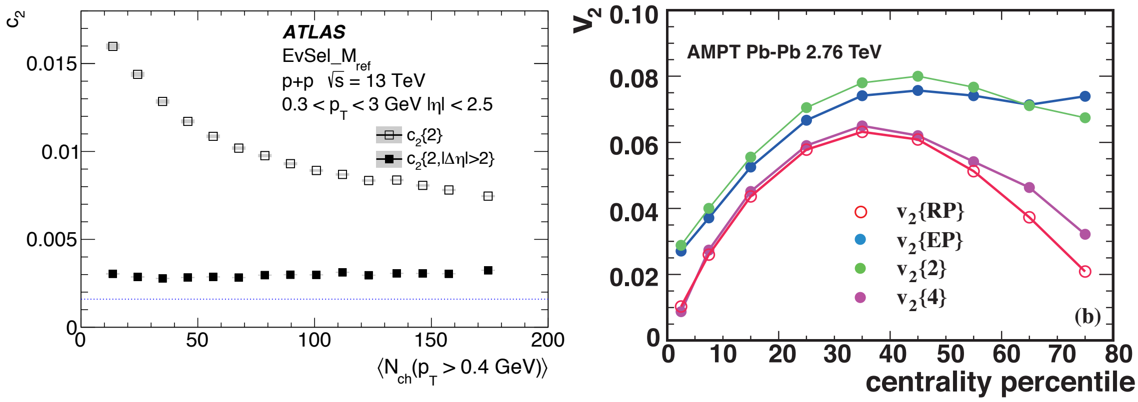

Figure 3.11: Left: $c_2\{2\}$ measured by ATLAS in p+p collisions at $\sqrt{s} = 13$ TeV with and without a rapidity gap. Figure extracted from [160]. Right: elliptic flow simulated in the AMPT framework [161] with different analysis techniques covered in the text. Figure extracted from [162].

arises from the two ways of pairing. The last possibility is that two particles are correlated due to flow and the other two are not that, with the corresponding combinatorial factor, gives $(2\times 2)v_n^2\delta_2$. The key point in this discussion is to substract Eqs. (3.34) and (3.35) in the following manner:

$$\begin{aligned} c_n\{4\} &\equiv \langle e^{(\mathrm{i}n(\phi_1+\phi_2-\phi_3-\phi_4))}\rangle - 2\langle e^{(\mathrm{i}n(\phi_1-\phi_2))}\rangle^2 \\ &= \langle v_n^4 + 4v_n^2\delta_2 + 2\delta_2^2 + \delta_4\rangle - 2\langle v_n^2 + \delta_2\rangle \\ &= \langle -v_n^4 + \delta_4\rangle \end{aligned} \tag{3.36}$$

where the last equality holds, if one neglects event-by-event flow and non-flow fluctuations i.e. $\langle v_2^2\rangle^2 = \langle v_2^4\rangle$. This procedure can be extended up to an arbitrary number of particles $c_n\{m\}$. Before explaining how to introduce these fluctuations note the remarkable feature of the four-particle cumulant: the δ_2 contribution has been removed and it only depends on δ_4. Hence, the main advantage of the cumulant analysis is that the contribution of non-flow correlations to v_n extracted from higher order cumulants are suppressed by powers of particle multiplicity. The flow coefficients from two and four particle cumulants are obtained from

$$v_n\{2\} = (c_n\{2\})^{1/2}, \;\; v_n\{4\} = (-c_n\{4\})^{1/4}. \tag{3.37}$$

Finally, if the value of v_n does not fluctuate and without non-flow $v_n = v_n\{2\} = v_n\{4\}$. If the event-by-event fluctuations of v_n are characterized by a standard deviation σ_{v_n}, and neglecting non-flow again, $v_n\{2\} = \sqrt{\langle v_n\rangle^2 + \sigma_{v_n}^2}$ [164]. When $\sigma_{v_n} \ll v_n$, $v_n\{4\} = \sqrt{\langle v_n\rangle^2 - \sigma_{v_n}^2}$ and the difference between $v_n\{2\}$ and $v_n\{4\}$ can be used to infer σ_{v_n} [165]. Due to all these arguments, an ordering of the flow harmonic coefficients $v_n\{2\} \geq v_n\{4\} \sim v_n\{6\}$ is expected. The differences between the discussed methods to compute v_n are nicely summarize in Fig. 3.11 using a simulation where the reaction plane is known by definition. Clearly, the estimator that provides a better agreement with $v_2\{\mathrm{RP}\}$ is $v_2\{4\}$.

The results for the elliptic flow in the three collision systems measured at the LHC are displayed in Fig. 3.12. Strikingly, a robust signal of a non-zero value of v_2 is extracted even in low multiplicity p+p collisions. Note that $v_2\{2\}$, even with the rapidity gap, is still higher due to residual non-flow contributions and fluctuations. These flow-like features were recently confirmed by the analyses of ATLAS and ALICE [160, 166]. Although in qualitative agreement, ATLAS and CMS results show non-negligible differences that affect both the absolute magnitude of v_n and their multiplicity dependence. The origin of these discrepancies has been identified as due to the method used to reduce non-flow contributions and the classification of the event activity. The different possibilities have been discussed both at a real data and a Monte Carlo level. For example, in [167, 160] the sensitivity of multi-particle azimuthal correlations to the multiplicity class definition in terms of kinematic variables such as $(\eta, p_{\rm T})$ and the multiplicity bin-width was studied. The non-negligible value of v_2 in p+p collisions have lead to an intense theoretical debate on the origin of these azimuthal correlations that will be the subject of the next section.

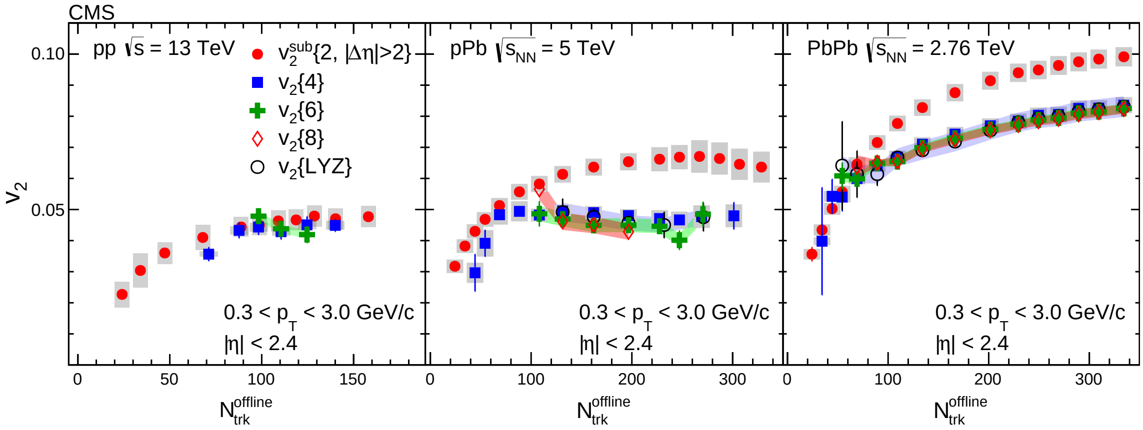

Figure 3.12: Integrated elliptic flow of unidentified charge particles as a function of the multiplicity. Figure extracted from [168].

The vast amount of high-precision data collected in the recent years at the LHC has allowed to go beyond the analysis of event-averaged observables and explore higher order moments of their probability distributions. More concretely, measurements of the correlations between different flow harmonics v_n have been performed. The so-called *symmetric cumulants* are defined as [159]

$$\begin{aligned}\mathrm{SC(n,m)} &\equiv \langle e^{\mathrm{i}(n\phi_1+m\phi_2-n\phi_3-m\phi_4)}\rangle - \langle e^{\mathrm{i}n(\phi_1-\phi_2)}\rangle\langle e^{\mathrm{i}m(\phi_1-\phi_2)}\rangle \\ &\sim \langle v_n^2 v_m^2\rangle - \langle v_n^2\rangle\langle v_m^2\rangle \end{aligned} \tag{3.38}$$

where the non-flow contributions have been neglected in the last step. Their normalized version

$$\mathrm{NSC(n,m)} \equiv \frac{\langle v_n^2 v_m^2\rangle - \langle v_n^2\rangle\langle v_m^2\rangle}{\langle v_n^2\rangle\langle v_m^2\rangle} \tag{3.39}$$

eliminates the dependence on the absolute magnitude of $v_{n(m)}$. The measurement of Eq. (3.38) would be zero by definition, if the fluctuations of v_n and v_m were totally uncorrelated, in the same way as the Pearson's correlation coefficient. Instead, a positive value of Eq. (3.38) implies that an event with $v_n > \langle v_n \rangle$ would be more likely to have $v_m > \langle v_m \rangle$ i.e. v_n and v_m are said to be correlated. Equivalently, if $\mathrm{SC(n,m)} < 0$ the two flow harmonic coefficients under consideration are anti-correlated. In particular, by measuring SC(2,3) we can gain information about initial state fluctuations whereas SC(2,4) is mostly sensitive to the strongly interacting medium properties [169].

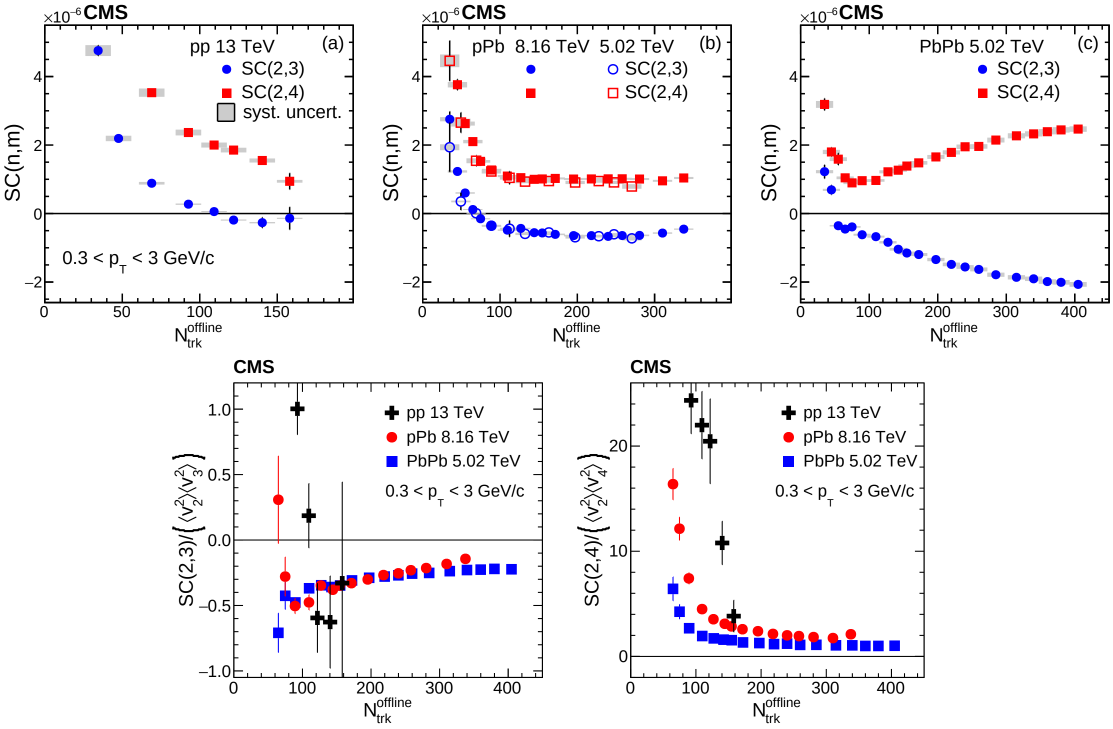

Figure 3.13: Top: SC(2,3) and SC(2,4) as a function of the multiplicity in the three collision systems. Bottom: normalized symmetric cumulants. Figures extracted from [170].

The experimental study of symmetric cumulants was pioneered by ALICE [171] in Pb+Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 2.76$ TeV. Recently, CMS has performed an experimental analysis of the symmetric cumulants as a function of the multiplicity in the three collision systems available at the LHC: p+p, p+Pb and Pb+Pb [170]. The experimental results, shown in Fig. 3.13, suggest a similar pattern across systems. Concretely, SC(2,4) is always positive although its multiplicity dependence varies from p+p to Pb+Pb. Further, NSC(2,4) is clearly modified when varying the system size. On the contrary, the sign of SC(2,3) is strongly multiplicity-dependent: at low multiplicities SC(2,3) is found to be positive. However, it turns out to be negative for very high multiplicities, $N_{\mathrm{trk}}^{\mathrm{offline}} > 60$ in both p+Pb and Pb+Pb and $N_{\mathrm{trk}}^{\mathrm{offline}} \sim 100$ in p+p. Moreover, NSC(2,3) in the high-multiplicity regime is found to have not only the same sign in the three collision systems but the same quantitative value as well.

Other measurement showing similarities between high-multiplicity proton-proton interactions and p+Pb and Pb+Pb is the enhanced production of multi-strange hadrons [172]. All together, the experimental data is constantly reigniting the debate on whether collective effects, precedently attributed to the formation of QGP droplets, are being observed in small collision systems such as p+p and p+Pb.

3.5.2 Theoretical perspective

From a theoretical point of view, we have seen that the well established paradigm of either Monte Carlo Glauber or CGC initial conditions followed by viscous hydrodynamic evolution has successfully described data on bulk observables in heavy-ion collisions. Turning to smaller systems the situation is less conclusive and two schools of thought have emerged to tackle the origin of these correlations.

- **Hydrodynamical evolution**: One option is to consider that there is no fundamental difference between the flow signals observed in p+p and Pb+Pb collisions and the *Panta Rhei* principle holds [173, 174]. This idea was adopted in [175] to simultaneously describe the integrated flow values in the three collision systems using viscous hydrodynamic evolution. A reasonable agreement with experimental data was found with the QGP parameters set to be the same in the three collision systems. Earlier, in the case of p+Pb interactions, a successful description of the harmonic flow was found in [176] using a similar set-up consisting on Glauber+event-by-event viscous relativistic hydrodynamics and with the EPOS model [177]. Regarding, the symmetric cumulants data a satisfactory description within the hydro paradigm was achieved in [169, 178] in the A+A case. The major criticism to this approach is related to the break down of a fluid description in a system that is neither in thermal equilibrium nor isotropic. The validity of hydrodynamics outside of equilibrium was discussed in [179]. Possibly, there is no way around quantitatively studying the Knudsen numbers reached in p+p collisions to obtain firm knowledge on the applicability of hydrodynamics.

- **Initial state dynamics**: Azimuthal anisotropies of multiparticle correlations can be generated in the Color Glass Condensate framework without the need of final state interactions due to QGP formation [180, 181]. This can be understood in a very simplified picture. For instance, consider a p+Pb collision in which uncorrelated partons from the proton scatter off the strong color fields of the nucleus. Then, each parton receives a transverse momentum kick in the scattering process. As we discussed in Section 3.3.1 a spatial extension, or domain, with a characteristic size of $1/Q_s$ can be attributed to the color fields. Therefore, when two partons, in the same color state, scatter off the same domain they acquire a similar transverse momentum kick that gives rise to a correlation. The diagrammatic representation of the scattering in coordinate space is displayed in Fig. 3.14. However, the diagrams contributing to the momentum correlations are suppressed by the number of color sources [180] and consequently the signal is weak. An additional degree of freedom, namely the orientation/polarization of the chromoelectric field that naturally collimate the partons, is included in the color domain model that enhances the flow signal [182]. Thus, two partons that were completely uncorrelated develop a momentum correlation that gives rise to a non-zero signal of v_n solely by means of the scattering process. In short, non-trivial initial state correlations can mimic collectivity. Remarkably, within this picture, a quantitative description of the ridge data was provided in [183]. A step forward has been recently presented in [184, 185] where, not only two but multi-particle azimuthal correlations and SC(n,m) are computed within the CGC in qualitative agreement with the p+Pb results. In the p+p case, the possibility of generating a non-zero v_n signal by converting the gluon distribution as obtained in classical Yang-Mills into hadrons with the Lund string fragmentation model contained in PYTHIA was explored in [186].

To finalize, the fact that both frameworks, conceptually orthogonal, are able to describe not all but some features of the experimental data keeps alive the flame of this exciting debate. Claims in favor of one particular approach should not be done solely on the basis of agreement with flow harmonic coefficients data. Such is a necessary but not sufficient condition. Together with stringent experimental predictions of the two models, a robust theoretical description across a wide kinematic range from p+p to A+A collisions seem unavoidable to work out which of them contains the most suitable dynamics underlying azimuthal correlations in small systems. An alternative and pragmatic view is to consider that this is not a black and white question. In that regard, a combination of both initial and final state momentum correlations cannot be discarded. The fundamental tasks in this gray area are to asses the strength of both effects in a given kinematic window and ensure a smooth matching between frameworks. This line of thought was explored in [187], on a weak coupling description of p+Pb interactions i.e. classical Yang-Mills+effective kinetic theory, finding that the relative importance of the initial state momentum correlations with respect to the ones built up in the evolution depend on the multiplicity of the event and the transverse momenta considered. In addition, a quantitative theory-to-data agreement in flow coefficients measurements in p+Pb has been achieved by feeding CGC-like initial momentum correlations into a viscous hydro evolution [188].

3.6 This thesis in the context of QGP physics

The parametrization of the geometry of the collision is mandatory in any theoretical model attempting to describe the striking experimental results that suggest collective behavior in small systems at the LHC, such as the non-zero value of the flow harmonic coefficients v_n. This information is indispensable regardless on whether the origin of these azimuthal correlations is attributed to the initial state dynamics or to the dynamical evolution of the expanding fireball as in the nucleus-nucleus case. Precisely the smaller systems exhibit a high degree of sensitivity to the description of the proton structure and its fluctuations. The origin of the fluctuations is intimately related with the quantum mechanical nature of the system. As an example, the importance of considering subnucleonic degrees of freedom when describing the elliptic flow in p+Pb within the IP-Glasma framework was

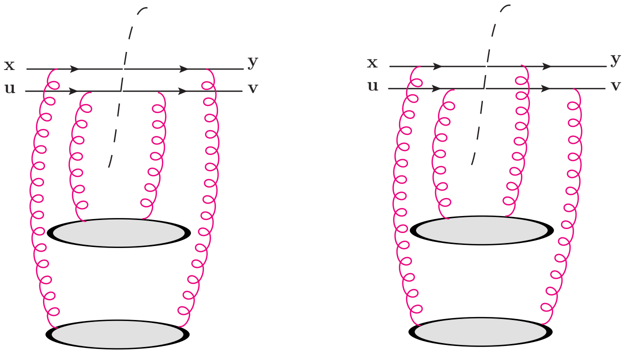

Figure 3.14: The grey blobs represent the domains of color fields in the target off which two independent quarks at transverse positions $(\mathbf{x},\mathbf{u})$ scatter. The momentum kick is equivalent, in the conjugate coordinate space, to a shift in the positions in the amplitude and conjugate amplitude. Left: independent scattering of the two quarks. Right: connected diagram that leads to genuine correlations. Figures adapted from [180].

realized in [188]. While current theoretical models differ in many aspects, as we have discussed in the previous section, all [188, 189, 190, 191, 192, 193] assume a similar picture of the proton that consists on:

- Three or more subnucleonic degrees of freedom dubbed valence quarks, constituent quarks or hot spots.
- Two relevant scales: the radii of the proton and the constituents.
- The subnucleonic components of each colliding proton are completely independent from each other i.e. they are uncorrelated.

However, in our view, it seems unrealistic that the components of the wave function of the proton do not know anything about each other. This claim is phenomenologically supported by the substantial impact of spatial correlation in the description of the hollowness effect (see Chapter 2). The subject of the next chapter is to go beyond the current assumptions on the proton structure in the literature and include spatial correlations among the sub-nucleonic degrees of freedom. For this purpose, a Monte Carlo Glauber event generator to study proton-proton interactions at different collision energies has been developed. Our work tackles the initial state of hadronic collisions i.e. $t=0$ in the language of Fig. 3.5.

Random name generator;
Wilco.

4

Initial state studies in p+p collisions

In this Chapter we investigate the effect of non-trivial spatial correlations between proton constituents, considered to be gluonic hot spots as introduced in Chapter 2, on the initial conditions of proton-proton collisions from ISR to LHC energies, i.e. $\sqrt{s} = 52.6, 7000, 13000$ GeV. Our analysis relies on a Monte Carlo Glauber approach including fluctuations in the hot spot positions and their entropy deposition in the transverse plane. We explore both the energy dependence and the effect of spatial correlations on the number of wounded hot spots, their spatial distribution and the eccentricities, ε_n, of the initial state geometry of the collision. A quantitative way to characterize the initial geometry anisotropy of the overlap region in a collision is to compute the spatial eccentricity moments (ε_n). A precise understanding of the geometry of the collision is the first step towards addressing the enigmatic flow signals in small collision systems at LHC energies. Precisely in a hydrodynamical picture, the flow harmonic coefficients are the response to the anisotropy in the initial density profile. Actually, a linear response, $v_n \propto \varepsilon_n$, was extensively assumed for a long time. Recently, non-linear terms have been shown to be non-vanishing [194, 195].

Moreover, we present a systematic study of the normalized symmetric cumulants, NSC(n,m), at the eccentricity level in proton-proton interactions at $\sqrt{s} = 13$ TeV. Further, we explore the dependence of our conclusions on the number of hot spots, the values of the hot spot radius and the repulsive core distance. Our results add evidence to the idea that considering spatial correlations between the subnucleonic degrees of freedom of the proton may have a strong impact on the initial state properties of high-energy proton-proton interactions.

This chapter is based on the publications:

- J. L. Albacete, H. Petersen, A. Soto-Ontoso, Correlated wounded hot spots in proton-proton interactions, Phys.Rev. C95 (2017) no.6, 064909 [196].

- J. L. Albacete, H. Petersen, A. Soto-Ontoso, Gluonic hot spots and spatial correlations inside the proton, Nucl.Phys. A967 (2017) 924-927 [197].

- J. L. Albacete, H. Petersen, A. Soto-Ontoso, Symmetric cumulants as a probe of the proton substructure at LHC energies, Phys.Lett. B778 (2018) 128-136 [198].

4.1 Setup

The proton is a quantum mechanical system. Therefore, in proton-proton collisions at the LHC its intrinsic properties, such as the arrangement of its constituents or how effective they are producing new particles, fluctuate on an event-by-event basis. Consequently, the model presented in Chapter 2 is not sufficient and has to be extended to accomodate these stochastic features. For that purpose, we developed a Monte Carlo Glauber event generator in ROOT [199] and C++. Similar codes for nucleus-nucleus collisions can be found in the literature [129, 200]. In this section, a detailed description of the building blocks is presented. In the following, for simplicity, a proton is considered to be formed by 3 hot spots. The impact of a varying number of hot spots $N_{hs} = (2, 4)$ is studied in Sec. 4.3 and 4.4.

4.1.1 Impact parameter distribution

First of all, the impact parameter of the collision is chosen randomly from the distribution

$$\mathrm{d}N_{\mathrm{ev}}/\mathrm{d}b \propto b \tag{4.1}$$

up to $b_{\max} = 2$ fm $\gtrsim 2R_p$ as displayed in the left side of Fig. 4.1. In our picture, the impact parameter is the distance between the centers of the two protons in the x-direction. Thus, the centers of the colliding protons are located at $(x, y) = (-b/2, 0)$ and $(b/2, 0)$. Furthermore, the z-component will be neglected in the whole calculation i.e. the analysis is performed exclusively in the transverse plane. We have checked that beyond $b = 2$ fm

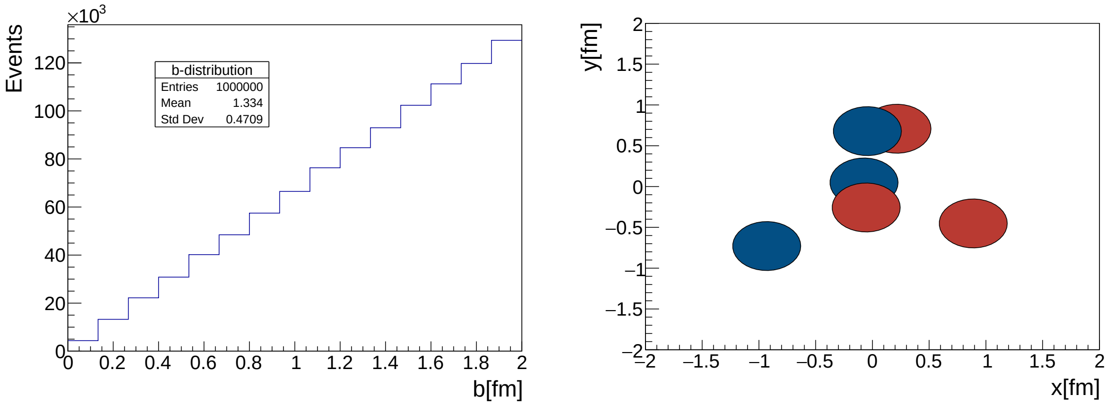

Figure 4.1: Left: impact parameter distribution as given by Eq. (4.1). Right: Typical event for p+p with the target constituents in red and the projectile ones in blue.

the number of events with at least one collision is negligible. The reaction plane is defined, in our case, by the (x, y)-plane.

4.1.2 Sampling hot spots positions

In each p+p event, after generating a random impact parameter for the collision, we sample the transverse positions of the three hot spots in each proton $\{\vec{s}_i\}$ according to the distribution defined in Chapter 2 i.e.:

$$D(\vec{s}_1, \vec{s}_2, \vec{s}_3) = C \prod_{i=1}^{3} e^{-s_i^2/R^2} \delta^{(2)}(\vec{s}_1 + \vec{s}_2 + \vec{s}_3) \times \prod_{\substack{i<j \\ i,j=1}}^{3} \left(1 - e^{-\mu|\vec{s}_i - \vec{s}_j|^2/R^2}\right). \tag{4.2}$$

It should be noted that the extension of Eq. (4.2) to an arbitrary number of hot spots is straight forward. As mentioned in the previous chapter, most of the models in the literature [201, 192, 193, 188, 175, 191, 202] implement a proton geometry following the two first terms of Eq. (4.2) i.e. the hot spots are distributed according to Gaussian functions with the natural constraint of fixing the centre of mass of the constituents system to the centre of the proton. However, with this set up the most probable configurations are the ones with three hot spots in the middle of the proton and the one with two hot spots fully or partially overlapping and the third one separated due to the δ-function like in the quark-diquark model. The third term of Eq. (4.2) allows us to go beyond these approaches by implementing short range repulsive correlations among all pairs of hot spots that effectively enlarge the mean transverse separation $|\vec{s}_i - \vec{s}_j|$ between them. It should be noted that we do not impose any kind of artificial minimum-distance between the hot spots as it is done in other works in the literature to mimic the short range correlations [200]. This would directly modify the single particle distribution. We generate the polar coordinates of the three hot spots, that are next easily converted into Cartesian ones, sampling $D(\vec{s}_1, \vec{s}_2, \vec{s}_3)$ for each proton. An illustrative example of a p+p interaction generated in this fashion is displayed on the right side of Fig. 4.1.

4.1.3 Collision criterion

Once the hot spots of projectile and target are located in the transverse plane, the probability of two hot spots to collide is sampled from the inelasticity density

$$G_{\text{in}}(d) = 2e^{-d^2/2R_{hs}^2} - (1 + \rho_{hs}^2)e^{-d^2/R_{hs}^2} \tag{4.3}$$

where d is the transverse distance between a pair of hot spots with radius R_{hs}, and ρ_{hs} is the ratio of real and imaginary parts of the hot spot-hot spot scattering amplitude. This collision probability results from a Gaussian parametrization of the elastic scattering amplitude (see Eq. (2.29)). We evaluate this probability for all pairs of hot spots and refer to them as *wounded* [203, 204], if they have suffered at least one collision. Thus, in each event, the maximum number of wounded hot spots N_w and collisions $N_{\rm coll}$ is 6 and 9 respectively. Another possibility that has been studied in the literature is to consider the number of binary collisions instead of the wounded hot spots scenario or a combination of both. We have tested that our main conclusions are not affected by this choice and take the wounded hot spot approach for simplicity. For each event, we keep track of the position of each wounded hot spot, (x_w, y_w), for later usage in the calculation of spatial distributions, eccentricities or any other quantity of interest.

4.1.4 Entropy deposition

Subsequently, we consider that each wounded hot spot located at (x_w, y_w) deposits a random amount of entropy smeared around the center of the wounded hot spot following a Gaussian prescription according to

$$s(x,y) = s_0 \frac{1}{\pi R_{hs}^2} \exp\left(-\frac{(x-x_w)^2+(y-y_w)^2}{R_{hs}^2}\right) \tag{4.4}$$

where s_0 fluctuates independently for each wounded hot spot. The choice of a Gaussian profile in Eq. (4.4) avoids unphysical spiked entropy deposition and endows our model with a more realistic description. To characterize the entropy deposition s_0 we rely on a very similar approach to the one recently proposed in [192]. Essentially, entropy deposition is directly related to the number of charged particles produced in p+p collisions. The charged hadron probability distribution in an incoherent description of the particle production process can be written as

$$\begin{aligned}\mathcal{P}(N_{\rm ch}) = \sum_{i=2}^{N_w} \mathcal{P}_w(i) \sum_{n_1,n_2,\ldots,n_i} & \mathcal{P}_{hw}(n_1)\mathcal{P}_{hw}(n_2)\ldots\mathcal{P}_{hw}(n_i) \\ & \times \delta(N_{\rm ch}-n_1-n_2-\ldots-n_i)\end{aligned} \tag{4.5}$$

where $\mathcal{P}_w$ is the probability distribution of i hot spots to be wounded and P_{hw} is the distribution of the number of hadrons produced by a single wounded hot spot. To make clearer the interpretation of Eq. (4.5) consider that we want to know the probability of producing $N_{\rm ch}=3$ in a collision where two hot spots have been wounded. Then, Eq. (4.5) reduces to:

$$\begin{aligned}\mathcal{P}(3) = \mathcal{P}_w(2)\Big[&\mathcal{P}_{hw1}(3)\mathcal{P}_{hw2}(0)+\mathcal{P}_{hw1}(0)\mathcal{P}_{hw2}(3) \\ &+\mathcal{P}_{hw1}(2)\mathcal{P}_{hw2}(1)+\mathcal{P}_{hw1}(1)\mathcal{P}_{hw2}(2)\Big]\end{aligned} \tag{4.6}$$

that is, the 3 charged particles can be produced by only one hot spot (the two first terms) or each of them contribute (the other two terms). According to Eq. (4.5) each wounded hot spot contributes independently to the total charged hadron multiplicity distribution in p+p collisions. An important comment is in order at this point. Particle production is treated incoherently, since we assume that each hot spot contributes the same way to the multiplicity distribution independent of the number of interactions. However, it seems reasonable to think that a certain degree of coherence should be included in a realistic model for particle production. By *coherence* we refer to the fact that an event where e.g. one hot spot in the projectile undergoes simultaneous scattering with three constituents in the target may not contribute in the same way to $\mathcal{P}(N_{\rm ch})$ as the incoherent superposition of three 1 vs. 1 interactions. For instance, in the IP-Glasma model the saturation scale is used as a degree of freedom to describe coherence [128]. Exploring a more realistic coherent description of the charged hadron multiplicity distribution is left for future work. Up to this point there is still one missing element in Eq. (4.5): the precise functional form for the hadron multiplicity distribution from each wounded hot spot $\mathcal{P}_{hw}(N_{\rm ch})$. The latest analysis of experimental data on charged hadron multiplicities by the LHC collaborations has revealed that a double negative binomial function provides a better description of the data than just a single one [205]. This is the choice adopted in this work to parametrize $\mathcal{P}_{hw}$:

$$\begin{aligned}\mathcal{P}_{hw}(N_{\rm ch}) = \alpha\frac{\Gamma(N_{\rm ch}+\kappa_1)\overline{n}_1^{N_{\rm ch}}\kappa_1^{\kappa_1}}{\Gamma(\kappa_1)N_{\rm ch}!(\overline{n}_1+\kappa_1)^{N_{\rm ch}+\kappa_1}}+\\ (1-\alpha)\frac{\Gamma(N_{\rm ch}+\kappa_2)\overline{n}_2^{N_{\rm ch}}\kappa_2^{\kappa_2}}{\Gamma(\kappa_2)N_{\rm ch}!(\overline{n}_2+\kappa_2)^{N_{\rm ch}+\kappa_2}}\end{aligned} \tag{4.7}$$

where $\Gamma(x)$ is the Euler Gamma function, the averages are given by $\overline{n}_i$, larger κ_i means smaller fluctuations and α is a mixing parameter. The parameters $\{\overline{n}_i,\kappa_i,\alpha\}$ are adjusted to reproduce the observed multiplicity distributions at all the collision energies considered in this work, independently. We achieve a good description of the data, $\chi^2/{\rm d.o.f} \sim 1.2{-}2$, for all cases as it is shown in the left side of Fig. 4.2. Nevertheless, there are small departures at low values of $N_{\rm ch}$ at $\sqrt{s}=13$ TeV and we overshoot the tail of the ISR data. We have not included in the fit the last 5 points of the experimental data of the ATLAS Collaboration at $\sqrt{s}=13$ TeV due to the large systematic uncertainties and their tiny contribution to the probability distribution.

Within the current context, the main purpose of an accurate description of the charged hadron multiplicity distribution is to provide phenomenological guidance to a non-measurable quantity, the shape of the entropy distribution i.e. s_0 in Eq. (4.4). In general, the negative binomial distribution can be expressed as a convolution of a Gamma and a Poisson distributions. Following the usual assumption that the particle emission is given by a Poissonian process with the mean proportional to the entropy deposited in the fluid element, the entropy distribution can be written as a double Gamma distribution:

$$\begin{aligned}\mathcal{P}(s_0) =&\alpha\frac{s_0^{\kappa_1-1}\kappa_1^{\kappa_1}}{\Gamma(\kappa_1)\overline{n}_1^{\kappa_1}}\exp\left(-\kappa_1 s_0/\overline{n}_1\right)+\\ &(1-\alpha)\frac{s_0^{\kappa_2-1}\kappa_2^{\kappa_2}}{\Gamma(\kappa_2)\overline{n}_2^{\kappa_2}}\exp\left(-\kappa_2 s_0/\overline{n}_2\right).\end{aligned} \tag{4.8}$$

The parameters $\{\overline{n}_i,\kappa_i,\alpha\}$ in Eq. (4.8) are identical to the ones of the negative binomial distribution, Eqs. (4.5-4.7), that yield a precise description of the measured multiplicity distributions (see Table 4.2).

4.1.5 Centrality classes

In this work we characterize the centrality of an event by its deposited entropy that is tightly related with the event multiplicity. In other words, the more entropy is deposited by the wounded hot spots, the more central the event. At this point it is worth to mention that the definition of *centrality* in small systems is a highly debated topic nowadays. Originally, the term *centrality* was borrowed from the heavy ion community. In these type of collisions a systematic correlation between the number of charged particles produced in an event (a measurable quantity) and the impact parameter (a theoretical variable) exists. Then, small values of $N_{\rm ch}$ are related to peripheral collisions i.e. large values of the impact parameter. In the case of small systems the correspondance between multiplicity and b becomes much weaker. These kind of systems are dominated by fluctuations that invalidate a geometric interpretation of multiplicity as a measure of collision *centrality* as was explicitly shown in [193] and later confirmed in our studies. A hard collision might produce lots of particles independently of b. To sum up, in our context by more central we mean more effective producing entropy and it does not always imply a small impact parameter. On the right side of Fig. 4.2 we show the integrated entropy deposition distribution for a particular choice of the parameters, the $r_c = 0.4$ fm case (to be described below), at $\sqrt{s} = 13$ TeV where S is computed in each event as

$$S = \sum_{i=2}^{N_w} s_0^i \tag{4.9}$$

and s_0 is given by Eq. (4.8). We have superimposed the division of the events in centrality classes depending on their contribution to the integrated entropy. Then, the $[0-1\%]$ bin contains one percent of the total number of events around the maximum value of the integrated entropy.

4.1.6 Parameters

Our model has four free parameters $\{R_{hs},\, R,\, r_c,\, \rho_{hs}\}$. For a given value of r_c we constrain $\{R_{hs},\, R,\, \rho_{hs}\}$ to reproduce the measured values of the total p+p cross section ($\sigma_{\rm tot}$) and the ratio of real and imaginary parts of the scattering amplitude (ρ) at each collision energy as was done in Sec. 2.4. Upon imposing these constraints we ensure that our results are phenomenologically compatible. Note that the values of the parameters of our model that fulfill the phenomenological conditions are not unique but rather conform a whole region of the parameter space. As shown below, we have chosen representative values of those allowed regions.

Regarding the correlation structure of the hot spots, Eq. (4.2), we have considered two extreme scenarios: the uncorrelated case labeled as $r_c = 0$ and a repulsive core of 0.4 fm labeled as $r_c = 0.4$. Being the main goal of this work to explore the net effect of correlations we have considered a third situation, $r_c = 0, nc$, in which we set the repulsive distance to 0 but choose the values of $\{R_{hs},\, R,\, \rho_{hs}\}$ as in the $r_c = 0.4$ case, not reproducing though the experimental values of $\sigma_{\rm tot}$ and ρ. However, although the rest of the parameters coincide, this is not enough to consider that differences between the results of $r_c = 0.4$ and $r_c = 0, nc$

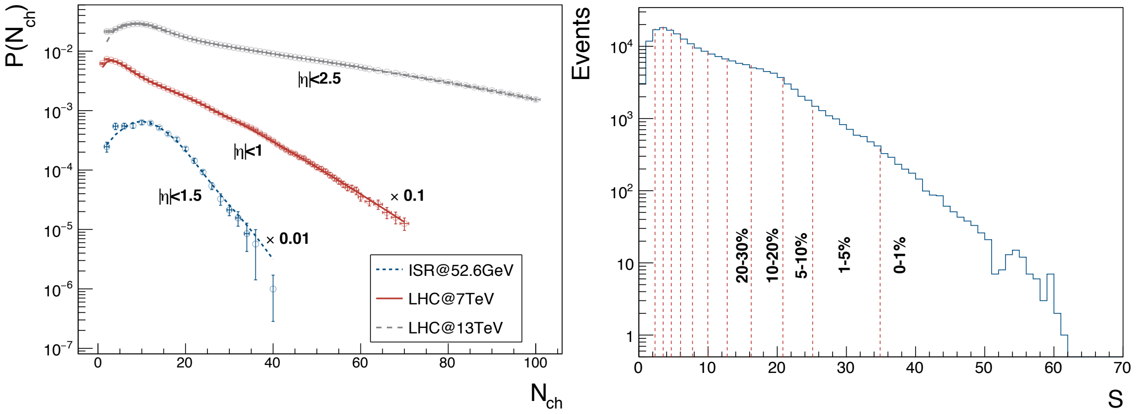

Figure 4.2: Left: Fit to the charged particle multiplicity distributions for different collision energies in the wounded hot spot model, Eqs. (4.5-4.7). The experimental data from top to bottom is taken from: ATLAS [206], ALICE [207] and ISR [208]. Note that each experiment has a different rapidity acceptance window, η, that influences the shape of the data. Also, the ALICE (red) and ISR (blue) curves are multiplied by 0.1 and 0.01 respectively. Right: Histogram of the integrated entropy deposition for the $r_c = 0.4$ fm case at $\sqrt{s} = 13$ TeV. Vertical red lines labelled by black numbers define centrality classes as fractions of the total number of events.

are then only attributable to the presence of short-range repulsive correlations. This is so because for the same values of $\{R_{hs},\, R,\, \rho_{hs}\}$ the hot spots of the correlated distribution have a larger mean transverse position (or r.m.s), $\langle s_1 \rangle$, defined as

$$\langle s_1 \rangle = \int s_1 \mathrm{d}\vec{s}_1 \mathrm{d}\vec{s}_2 \mathrm{d}\vec{s}_3 D(\vec{s}_1, \vec{s}_2, \vec{s}_3) \tag{4.10}$$

where $D(\vec{s}_1, \vec{s}_2, \vec{s}_3)$ is given by Eq. (4.2), than for $r_c = 0$. In order to avoid this artificial swelling we have included one last scenario, labeled as "$\langle s_1 \rangle$ fixed", in which $\{R_{hs},\, \rho_{hs}\}$ are the same as in the $r_c = 0.4$ case but R is chosen to reproduce the $\langle s_1 \rangle$ of the correlated distribution.

The values of $\{R_{hs},\, R,\, r_c,\, \rho_{hs}\}$ are shown in Table 4.1. Further, the parameters of the double Gamma distribution that controls the entropy deposition s_0 obtained after fitting the multiplicity distributions are given in Table 4.2. Note that the fits shown in Fig. 4.2 correspond to the $r_c = 0.4$ case. Good quality fits were found in the other scenarios as well.

Once the building blocks of our model have been presented in the next sections we display its results.

4.2 Basic quantities in a Monte Carlo Glauber calculation

Before addressing the two relevant observables for flow studies i.e. spatial eccentricities and their fluctuations in terms of symmetric cumulants, it is interesting to explore the effect of spatial correlations in core elements of any Monte Carlo Glauber simulation such as the average number of wounded objects or the radial distribution of the participants. The following results have been obtained after generating 500k events with the set of parameters corresponding to $\sqrt{s}=7$ TeV. The results are qualitatively the same for ISR and LHC-Run II energies. The averages have been performed over the number of events with at least one hot spot-hot spot collision.

	$r_c = 0.4$ fm		$r_c = 0$		$\langle s_1 \rangle$ fixed
$\sqrt{s}$ [GeV]	R_{hs} [fm]	R [fm]	R_{hs} [fm]	R [fm]	R [fm]
52.6	0.19	0.72	0.23	0.73	0.84
7000	0.3	0.69	0.39	0.65	0.83
13000	0.32	0.76	0.41	0.75	0.87

Table 4.1: Parameters of the hot spot distribution and the inelasticity profile in Eqs.(4.2-4.3) for various p+p collision energies with ($r_c = 0.4$ fm) and without ($r_c = 0$) short-range repulsive correlations. We set $\rho_{hs} = 0.1$ in all cases. On the last column, the values of R for the "$\langle s_1 \rangle$ fixed" case are shown.

$\sqrt{\mathbf{s}} = \mathbf{62.5}$ **GeV**	$\overline{n}_1$	κ_1	$\overline{n}_2$	κ_2	α
$r_c = 0.4$ fm	7.81	36.10	2.46	2.20	0.52
$r_c = 0$	5.61	49.12	0.07	0.03	0.77
$r_c = 0, nc$	6.48	26.58	0.44	41.61	0.72
$\langle s_1 \rangle$ fixed	6.70	18.14	0.43	27.98	0.73

$\sqrt{\mathbf{s}} = \mathbf{7}$ **TeV**	$\overline{n}_1$	κ_1	$\overline{n}_2$	κ_2	α
$r_c = 0.4$ fm	11.86	2.45	2.04	1.49	0.27
$r_c = 0$	10.86	2.27	1.72	1.68	0.24
$r_c = 0, nc$	11.47	2.40	1.92	1.54	0.26
$\langle s_1 \rangle$ fixed	11.69	2.40	1.98	1.50	0.27

$\sqrt{\mathbf{s}} = \mathbf{13}$ **TeV**	$\overline{n}_1$	κ_1	$\overline{n}_2$	κ_2	α
$r_c = 0.4$ fm	27.29	1.51	4.68	1.66	0.37
$r_c = 0$	26.26	1.55	4.16	1.79	0.31
$r_c = 0, nc$	21.70	1.51	4.38	2.50	0.48
$\langle s_1 \rangle$ fixed	26.47	1.42	4.54	1.73	0.37

Table 4.2: Default values of the parameters of the double Gamma distribution that characterizes the fluctuating amount of entropy each wounded hot spot deposits, s_0, for $\sqrt{s} = 52.6, 7000, 13000$ GeV. The errors of the parameters arising in the fitting procedure are $\sim \pm 0.01$.

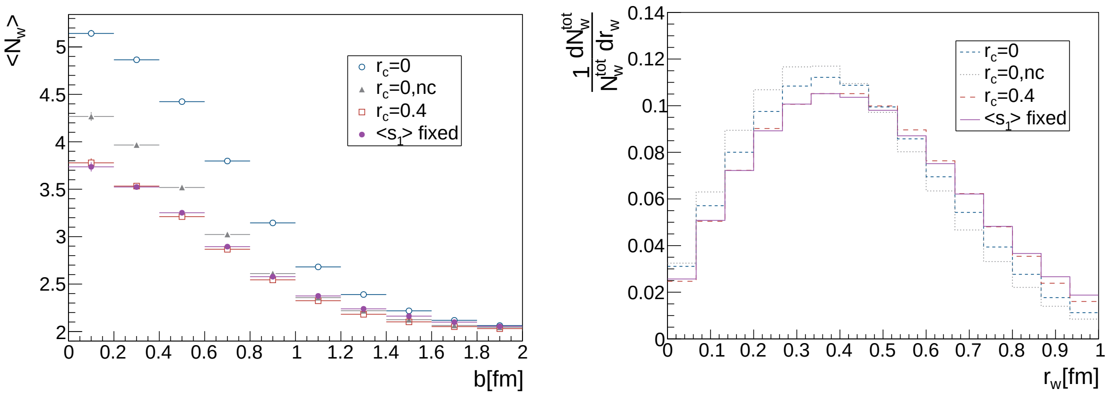

Figure 4.3: Left: Average number of wounded hot spots for different impact parameter bins of the collision. The horizontal lines indicate the width of the bins. Right: Normalized radial distribution of the wounded hot spots with respect to the reaction plane.

4.2.1 $\langle N_w \rangle$ vs. b

We begin our analysis by computing the average number of wounded hot spots in p+p collisions as a function of the impact parameter b for the four different scenarios introduced above. The results are shown in the left part of Fig. 4.3. We note that the qualitative behavior of the impact parameter dependence of $\langle N_w \rangle$ is not affected by the inclusion of correlations. For instance, the number of wounded hot spots is larger in central collisions ($b=0$) than in peripheral ones, as expected. However, in central to moderately peripheral collisions, $0<b<0.8$ fm, the average number of wounded hot spots is smaller in the correlated scenario (squares vs. empty dots/triangles in Fig. 4.3). We have also computed the mean number of wounded hot spots in a proton-proton interaction defined as

$$\overline{N}_w = (\sum_{i=1}^{N_{\rm ev}} N_w^i)/N_{\rm ev} \tag{4.11}$$

where $N_{\rm ev}$ is the total number of events with at least one collision. We find that $\overline{N}_w$ is slightly reduced $\sim 5\%$ in the $r_c = 0.4$ case with respect to fixing $\langle s_1 \rangle$. Then, a very basic element of all Monte Carlo Glauber calculations i.e. the mean number of wounded objects, hot spots in our case, is already affected by the modification of the initial geometry of the collision as it is illustrated in Fig. 4.3. Nevertheless, the dominant scale affecting $\langle N_w \rangle$ is not r_c but R_{hs} i.e. the bigger the hot spots, the larger the value of $\langle N_w \rangle$. It is worth to mention that, in agreement with [192], we find moderate variations in the mean number of wounded hot spots as the energy increases. Indeed, for the $r_c = 0.4$ case we obtain $\overline{N}_w = 2.3,\ 2.74,\ 2.75$ at $\sqrt{s} = 52.6$ GeV, 7 TeV and 13 TeV respectively. The rising behavior of $\overline{N}_w$ with increasing collision energy can be directly attributed to the growth of R_{hs} as depicted in Table 4.1. In other words, in a geometrical framework as it is the Monte Carlo Glauber model, bigger hot spots translate into more collisions between them.

4.2.2 Radial distribution of wounded hot spots

Another important feature of the presence of repulsive correlations is their effect on the spatial distribution of the hot spots. In the right part of Fig. 4.3 we compare the normalized radial distribution of the wounded hot spots, characterized by their polar coordinate $r_w = \sqrt{x_w^2 + y_w^2}$, resulting from the uncorrelated and correlated scenarios. One sees that for $r_c = 0.4$ fm, the radial distribution gets broader and its mean value is shifted to larger values than in the uncorrelated case. Therefore, a plausible interpretation in a geometrical picture is that when including repulsive correlations the probability to find wounded hot spots on the edges of the interaction region is increased.

4.2.3 $\langle S \rangle$ vs. b and $\mathcal{P}_w(i)$

The behavior of the average entropy deposited, $\langle S \rangle$, as a function of the impact parameter is depicted in Fig. 4.4. We found a similar trend in the four cases considered in this work: $\langle S \rangle$ decreases as a function of the impact parameter as it should be by construction in our model i.e. on average the more central the smaller the impact parameter or, equivalently, larger number of wounded hot spots, the larger the entropy deposited. It is important to emphasize that this is true on average and not on an event-by-event basis as we have already discussed. We see that the difference between "$r_c = 0.4$" and "$\langle s_1 \rangle$ fixed" cases is small in Fig. 4.4.

An easy to way to crosscheck if the parameters of the multiplicity distribution extracted from the fits are in agreement with other features of the Monte Carlo calculation is to compute the average probability of i hot spots to be wounded. This factor, $\mathcal{P}_w(i)$ takes part in the functional shape of $\mathcal{P}(N_{\mathrm{ch}})$ (see Eq. (4.5)) whose parameters are fitted to data. From Fig. 4.4 we observe how correlations favor collisions with a small number of wounded hot spots. This has a direct consequence: in order to reproduce the probability distribution for charged particles each wounded hot spot has to be more efficient or, equivalently produce more particles on average than in the uncorrelated cases. In other words, when trying to describe $\mathcal{P}(N_{\mathrm{ch}})$ a trade-off between the mean number of wounded hot spots and the average number of produced particles by hot spot exists. That is, if the mean number of wounded hot spots decreases from the correlated to the uncorrelated case it has to be compensated by an increase on the averages in the negative binomial ($\overline{n}_1, \overline{n}_2$). In fact, this argument is confirmed by the parameters extracted from the fits displayed in Table. 4.2.

4.3 Spatial eccentricity moments

A quantitative measurement of the initial anisotropy of the geometry in a collision, which results into an asymmetric transverse flow by means of the pressure gradients, is given by the spatial eccentricities that are defined as

$$\varepsilon_n = \frac{\sqrt{\langle r^n \cos(n\phi) \rangle^2 + \langle r^n \sin(n\phi) \rangle^2}}{\langle r^n \rangle} \tag{4.12}$$

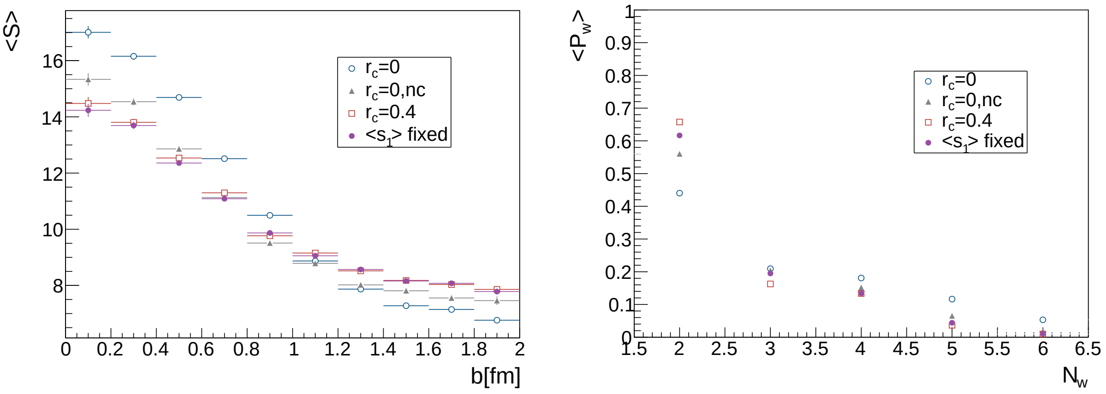

Figure 4.4: Left: Average entropy deposited as a function of the impact parameter. Right: Average probability of i hot spots to be wounded that enters into Eq. (4.5).

where $\langle\cdot\rangle$ denotes the average weighted by the entropy deposition i.e. $\int \mathrm{d}x\mathrm{d}y(\cdot)s(x,y)$ and includes a sum over all wounded hot spots (see Eq. (4.4)). By construction, $0 \leq \varepsilon_n \leq 1$. For $n=2$, Eq. (4.12) reduces to the standard definition of the eccentricity of an ellipse i.e. $\varepsilon_2 = 0$ for a circle. In our calculation ε_n are defined on an event-by-event basis with respect to the participant plane as described in the next section.

4.3.1 Participant plane

In a Glauber Monte Carlo event, when calculating eccentricities of a system that suffers fluctuations, in our case of the hot spots positions and of their entropy deposition, one important aspect should be noted. To make the discussion more intuitive we take ε_2 as an example but the idea generalizes to any moment. In a given event, the minor axis of the ellipse in the transverse plane formed by the wounded hot spots does not necessarily point along the impact parameter direction as given in Fig. 4.1. In other words, the principle axes of the ellipse may be tilted by an angle ψ_2 with respect to the (x,y)-directions. This new reference frame defined by the tilted axes is called *participant plane*. In order to avoid that the spatial eccentricities lose their power of characterizing the anisotropy of the collision one has to transform the wounded hot spots coordinates to ensure that the minor and major axes of their distribution align with the x and y axes. This situation is illustrated in Fig. 4.5 where the green circles represent the (x_w, y_w) positions of the wounded hot spots and the purple ones correspond to the transformed coordinates. This transformation is a two-step procedure.

First, Eq. (4.12) is valid whenever $\langle x\rangle = \langle y\rangle = 0$. Then, the coordinates are shifted such that $(0,0)$ coincides with the center of mass of the participants system. In our case, the wounded hot spots create an entropy distribution in the transverse plane as given by Eq. (4.4). Therefore, the shifted distribution, $s'(x,y)$, is given by

$$s'(x,y) = s\left(x - \int xs(x,y)\mathrm{d}x\mathrm{d}y, y - \int ys(x,y)\mathrm{d}x\mathrm{d}y\right). \tag{4.13}$$

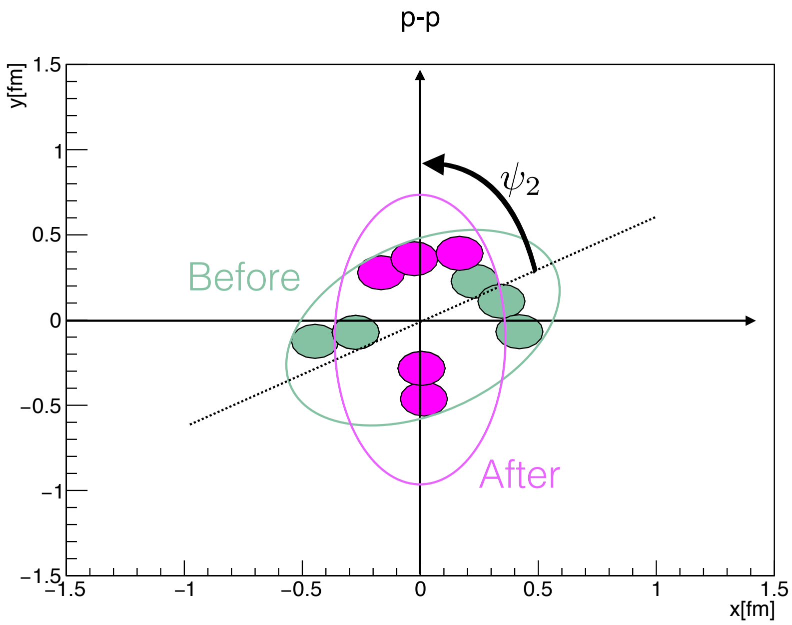

Figure 4.5: Illustrative example of a $\sqrt{s} = 7$ TeV p+p event with 5 wounded hot spots (green disks). The principal axes of the area formed by the wounded hot spots are tilted by an angle ψ_2 with respect to the reaction plane given by the axes (x and y) of the transverse plane. The purple disks represent the transformed coordinates.

Next we determine the angular orientation of the ε_n plane from

$$\psi_n = \frac{1}{n}\arctan 2\left(\frac{\int \mathrm{d}x\mathrm{d}y(x^2+y^2)^{n/2}\sin(n\phi)s'(x,y)}{\int \mathrm{d}x\mathrm{d}y(x^2+y^2)^{n/2}\cos(n\phi)s'(x,y)}\right). \tag{4.14}$$

Then, the participant spatial eccentricities are defined on an event-wise basis as:

$$\varepsilon_n = \frac{\int \mathrm{d}x\mathrm{d}y(x^2+y^2)^{n/2}\cos(n(\phi-\psi_n))s'(x,y)}{\int \mathrm{d}x\mathrm{d}y(x^2+y^2)^{n/2}s'(x,y)} \tag{4.15}$$

A good test to check if the method has been correctly implemented consists in simplifying Eq. (4.15) by removing the smearing effect and the entropy fluctuations. In this case, $x'^i = x^i_w - x^{\mathrm{C.o.M}}_w$,$y'^i = y^i_w - y^{\mathrm{C.o.M}}_w$ and the ψ_n angle is given by

$$\psi_n = \frac{1}{n}\arctan 2\left(\frac{\sum\limits_{i=1}^{N_w} r'^i\sin(n\phi'^i)}{\sum\limits_{i=1}^{N_w} r'^i\cos(n\phi'^i)}\right). \tag{4.16}$$

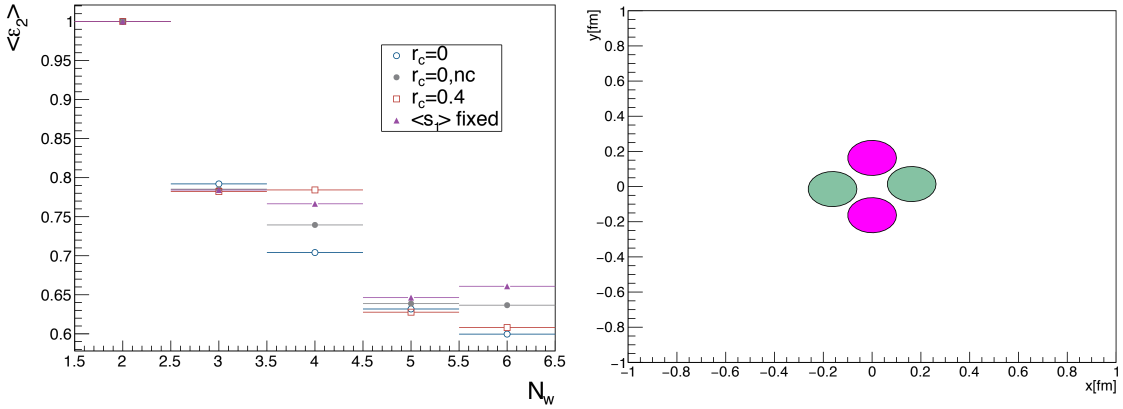

Figure 4.6: Left: Average values of the eccentricity, ε_2, for $r_c = 0$ (blue empty circle), $r_c = 0.4$ fm (red empty square), $r_c = 0, nc$ (grey filled circle) and $\langle s_1 \rangle$ fixed (purple filled circle) as a function of the number of wounded hot spots. Right: Example of an event with two wounded hot spots before (green disks) and after (purple disks) rotating to the participant plane.

Therefore, the computation of the spatial eccentricity moments is given by

$$\varepsilon_n = \frac{\sqrt{\sum_{i=1}^{N_w}\left((r^i_{pp})^n \cos(n\phi^i_{pp})\right)^2 + \sum_{i=1}^{N_w}\left((r^i_{pp})^n \sin(n\phi^i_{pp})\right)^2}}{\sum_{i=1}^{N_w}(r^i_{pp})^n} \tag{4.17}$$

where (r^i_{pp}, ϕ^i_{pp}) are the polar transformation of $(x^i_{pp} = x'^i \cos\psi_n - y'^i \sin\psi_n,\ y^i_{pp} = y'^i \cos\psi_n + x'^i \sin\psi_n)$. In Fig. 4.6 the event-averaged value of ε_2 as a function of the number of wounded hot spots is shown. The fact that $\langle\varepsilon_2\rangle$ reaches unity for $N_w = 1$ is a direct consequence of the participant plane method. When only two hot spots are wounded, they sit symmetrically on the vertical axes after rotating by ψ_2 as depicted on the right side of Fig. 4.6. In this case, $\langle\varepsilon_2\rangle = 1$ (and $\varepsilon_3 = 0$) by definition. This artifact is washed out once the smearing and entropy fluctuations are taken into account.

After cross-checking the implementation of the participant plane calculation, in the following we compute ε_n as given by Eq. (4.15) and refer to them simply as *spatial eccentricities*. We present our results in two different cases: all the events are selected (minimum bias) and only the events on the $[0-1\%]$ centrality class (ultra-central collisions) as defined in Fig. 4.2 are considered.

4.3.2 Minimum bias

The probability distribution of ε_2 is shown in Fig. 4.7: the eccentricity is reduced in the correlated scenario compared to the rest of the cases including the one with $\langle s_1 \rangle$ fixed. We hence conclude that the probability of having smaller values of the eccentricity in

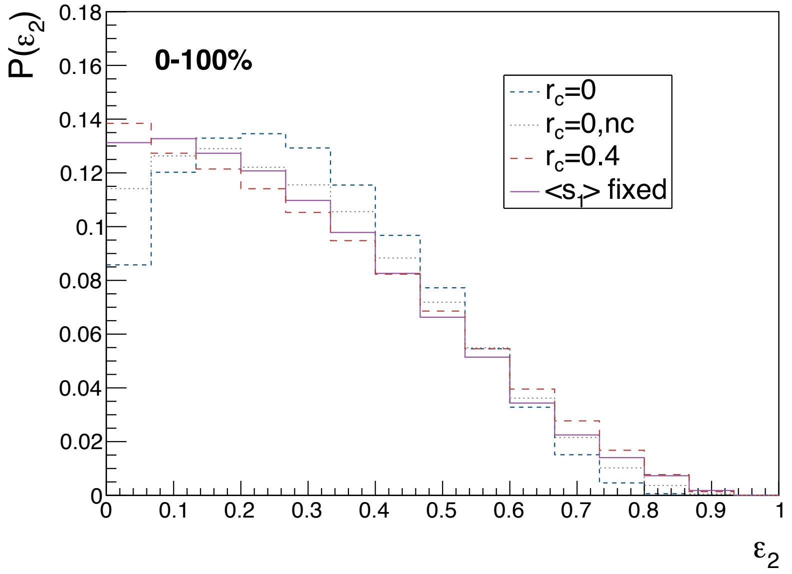

Figure 4.7: Probability distribution of the eccentricity, ε_2, for $r_c = 0$ (blue short-dashed line), $r_c = 0.4$ fm (red long-dashed line), $r_c = 0, nc$ (grey dotted line) and $\langle s_1 \rangle$ fixed (purple solid line).

a proton-proton interaction is increased when repulsive short-range correlations are included. In essence the eccentricity is a direct measurement of the anisotropy of the interaction region between the x and y directions. Thus, the results presented in Fig. 4.7 suggest that the characteristic ellipsoidal shape of the interaction region between the two protons is replaced by a more round one (with smaller eccentricity) in the correlated scenario. However, all cases exhibit a broad probability distribution of ε_2 due to the highly fluctuating nature of the system.

In addition, the effect of correlations between the constituents of the proton is shown to be qualitatively the same as in the nucleus case [209] but has a stronger impact on the numerical values of ε_2. This is not a surprising result as smaller systems are expected to be more sensitive to the fine details of the geometry than the complex A+A case, where the net effect of these subtleties is washed out by the accumulation of uncorrelated nucleon-nucleon collisions. However, this study being a multiparametric one the magnitude of the eccentricity's depletion could vary depending on the values of $\{R_{hs}$, R, r_c, $\rho_{hs}\}$ provided that it will always decrease when including correlations among subnucleonic degrees of freedom in the proton.

Further, we show in Fig. 4.8 the probability distribution for the triangularity, $\mathcal{P}(\varepsilon_3)$. The origin of odd eccentricity moments, in our model, is not driven by the geometry of the collision but rather by the fluctuations in both the entropy deposition and the positions of the hot spots. The main reason is that our spatial distributions are symmetric with respect to the y-axis, $\langle y \rangle = 0$, so in absence of fluctuations all the odd eccentricity moments would exactly vanish. Compared to ε_2, we observe how the role of spatial correlations for ε_3

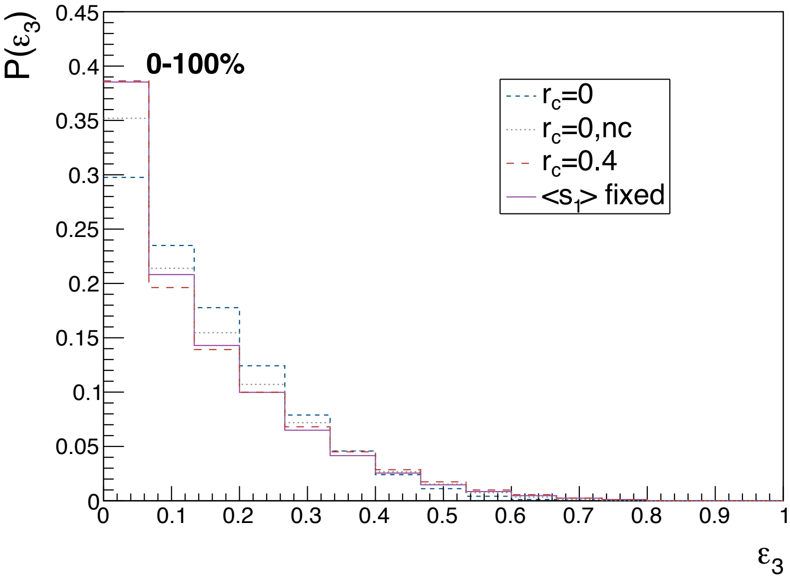

Figure 4.8: Probability distribution of the triangularity, ε_3, for $r_c = 0$ (blue short-dashed line), $r_c = 0.4$ fm (red long-dashed line), $r_c = 0, nc$ (grey dotted line) and $\langle s_1 \rangle$ fixed (purple solid line).

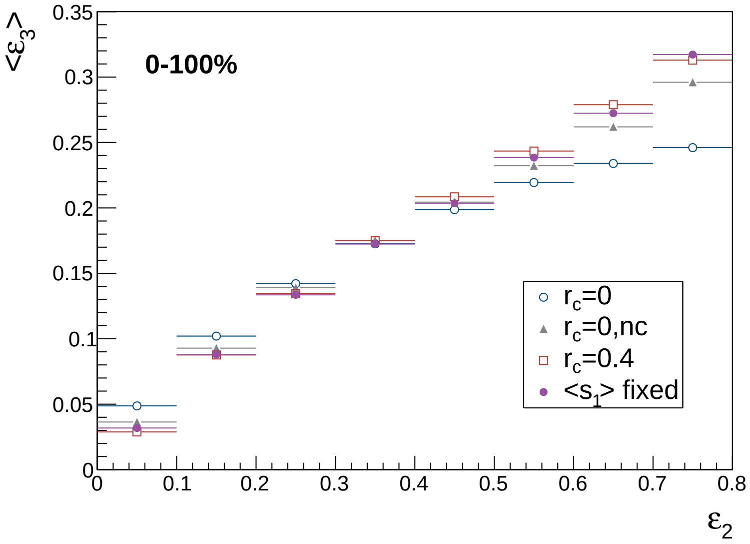

Figure 4.9: Average value of the triangularity as a function of the eccentricity for $r_c = 0$ (blue empty circle), $r_c = 0.4$ fm (red empty square), $r_c = 0, nc$ (grey filled circle) and $\langle s_1 \rangle$ fixed (purple filled circle).

is weaker. Nevertheless, the triangularity of the proton-proton interaction in our model shows the same qualitative behavior as the eccentricity i.e. is reduced in the correlated scenario.

To conclude this Section the correlation between ε_3 and ε_2 is displayed in Fig. 4.9. In agreement with other works in the literature [193], the value of ε_3 is always smaller than ε_2 and the correlation among them is positive. Note that, the impact of spatial correlations is the opposite depending in which regime of values of ε_2 we are. That is, for $\varepsilon_2 \geq 0.4$, repulsive correlations generate a more triangular interaction region while for smaller values of ε_2 the conclusion is reverted. A similar situation will be discussed in the next section. Nevertheless, by comparing the $\varepsilon_3 - \varepsilon_2$ correlation between the $\langle s_1 \rangle$ and $r_c = 0.4$ fm cases we conclude that it is a rather insensitive observable to the presence of spatial correlations.

4.3.3 Ultra-central collisions

Motivated by the phenomenological interest in very high-multiplicity p+p collisions we have computed the probability distributions $\mathcal{P}(\varepsilon_{2(3)})$ for the $0-1\%$ centrality class i.e. after imposing a cut on the entropy deposition S as given by Fig. 4.2. Two remarkable results can be extracted by comparing Figs. 4.7-4.8 (minimum bias) with Fig. 4.10 (ultra-central collisions). First, we observe how the probability distributions $\mathcal{P}(\varepsilon_{2(3)})$ are shifted towards larger values when selecting ultra-central events both in the uncorrelated and correlated scenarios. Next, focusing on the role of correlated constituents in this high-entropy context we observe that it turns out to favor higher values of ε_2 and ε_3 when compared to the uncorrelated scenario. Thus, we find that the consequence of having correlated constituents inside the proton is the opposite in ultra-central collisions than in minimum bias. To sum up, in the $0-1\%$ centrality class the net effect of correlations is to increase the probability of having larger values of $\varepsilon_{2(3)}$ whereas in the minimum bias case this probability is diminished.

The difference on the effect of spatial correlations between the minimum bias case and the ultra-central can be neatly deduced from Figs. 4.11-4.12. We represent the average values of $\varepsilon_{2(3)}$ for different centrality classes. A common trend is observed both in the correlated and uncorrelated cases: while $\langle \varepsilon_{2(3)} \rangle$ is barely centrality independent in the mid-central to peripheral collisions it increases significantly in the very central region i.e. in the events with higher entropy deposition. Regarding the effect of spatial correlations we notice that they increase $\langle \varepsilon_{2(3)} \rangle$ for the higher entropic events with respect to the uncorrelated cases as we have already shown in Fig. 4.10. Minimum bias collisions are not dominated by these infrequent extremely entropic events but by the ones with a smaller entropy production. In this case, we see in Figs. 4.11-4.12 how the net effect of correlations in the peripheral/less entropic bins is to reduce $\langle \varepsilon_{2(3)} \rangle$. Furthermore, the quantitative difference in $\langle \varepsilon_{2(3)} \rangle$ between the correlated and uncorrelated scenarios is larger in the ultra-central events than in minimum bias. Thus, we conclude that the net effect of correlated constituents is larger in ultra-central collisions.

Certainly, it would be desirable to obtain a quantitatively precise description of experimental data on v_n. Unfortunately, this is not possible within the current status of our

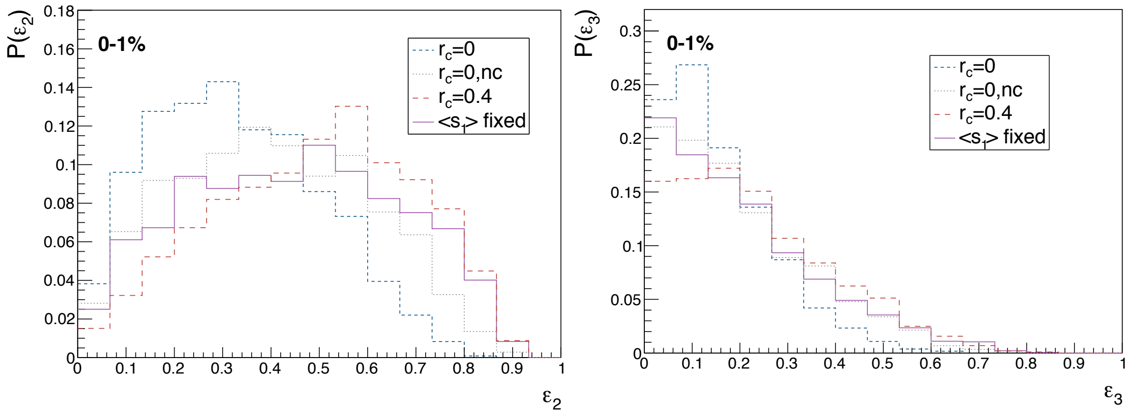

Figure 4.10: Average values of the eccentricity, ε_2, (left) and the triangularity, ε_3, (right) for $r_c = 0$ (blue empty circle), $r_c = 0.4$ fm (red empty square), $r_c = 0, nc$ (grey filled circle) and $\langle s_1 \rangle$ fixed (purple filled circle) as a function of the number of wounded hot spots.

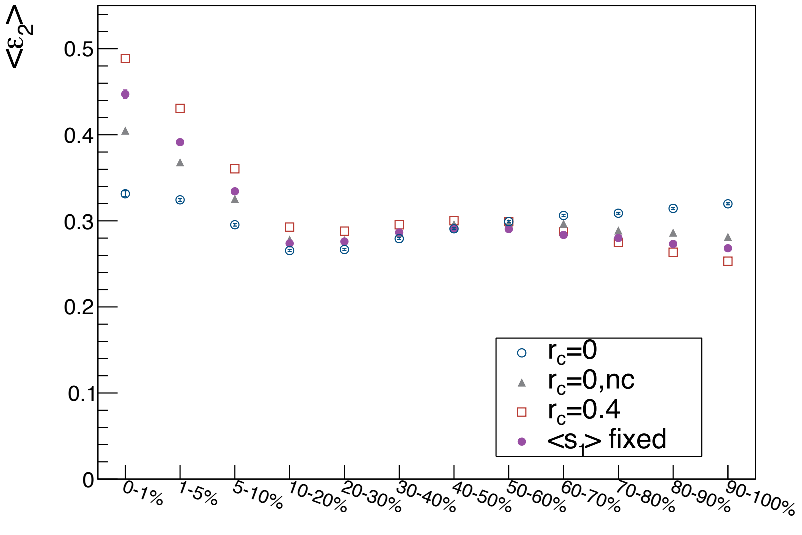

Figure 4.11: Average values of the eccentricity, ε_2, for $r_c = 0$ (blue empty circle), $r_c = 0.4$ fm (red empty square), $r_c = 0, nc$ (grey filled triangle) and $\langle s_1 \rangle$ fixed (purple filled circle) as a function of the centrality range.

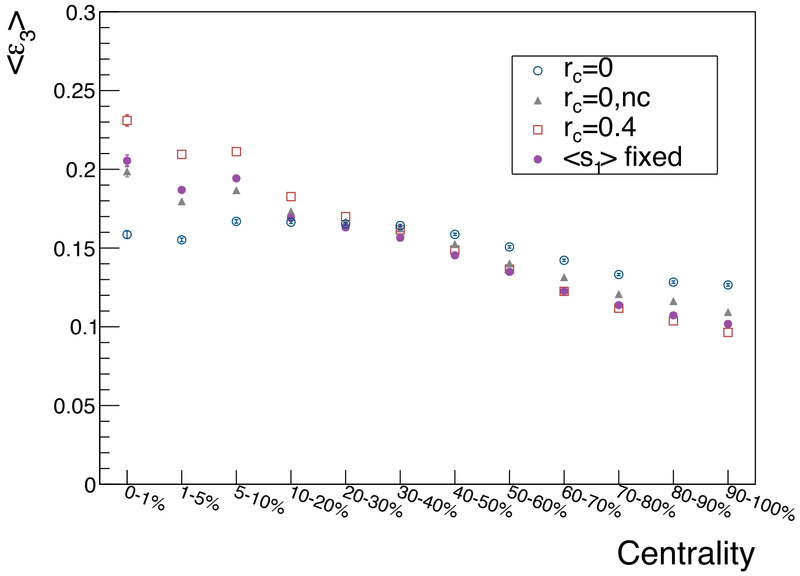

Figure 4.12: Average values of the triangularity, ε_3, for $r_c = 0$ (blue empty circle), $r_c = 0.4$ fm (red empty square), $r_c = 0, nc$ (grey filled triangle) and $\langle s_1 \rangle$ fixed (purple filled circle) as a function of the centrality range.

relatively simple geometric model. Our calculation lacks of any kind of information on the momenta of the produced particles. Thus, we consider that only after feeding our model into a hydrodynamic evolution, currently a work in progress, a full comparison with data would be meaningful and realistic. Nevertheless, one can grasp a feeling by computing ratios of cumulants, as defined in Eq. (3.37), because in first approximation they are independent of the hydrodynamic response [210] i.e.:

$$\frac{v_n\{\mu\}}{v_n\{\nu\}} = \frac{\varepsilon_n\{\mu\}}{\varepsilon_n\{\nu\}}. \tag{4.18}$$

In the case of the eccentricity,

$$\varepsilon_2\{2\} = \langle \varepsilon_2^2 \rangle, \ \ \varepsilon_2\{4\} = (2\langle \varepsilon_2^2 \rangle^2 - \langle \varepsilon_2^4 \rangle)^{1/4} \tag{4.19}$$

The equality in Eq. (4.18) only holds if the scaling between eccentricity and flow is linear. Although deviations from this linear response are expected [194], it is interesting to check whether our model can capture the general trends of the data. The comparison between eccentricity and flow fluctuations is displayed in Fig. 4.13. Note that the experimental data is given in terms of the flow harmonic coefficients as a function of the $\mathrm{N}_{\mathrm{trk}}^{\mathrm{offline}}$. This variable contain the number of reconstructed charged particles after imposing p_{T} and η cuts which vary from one experiment to the other. Therefore, we refuse to present the results of our model overlaid on top of the experimental data as the x-axes, although sharing the same spirit of representing particle/entropy production, do not match i.e. it will not be an apples-to-apples comparison. Presenting our results as a function of N_{ch} as shown in Fig. 4.2 will not help to directly compare to the experimental measurement

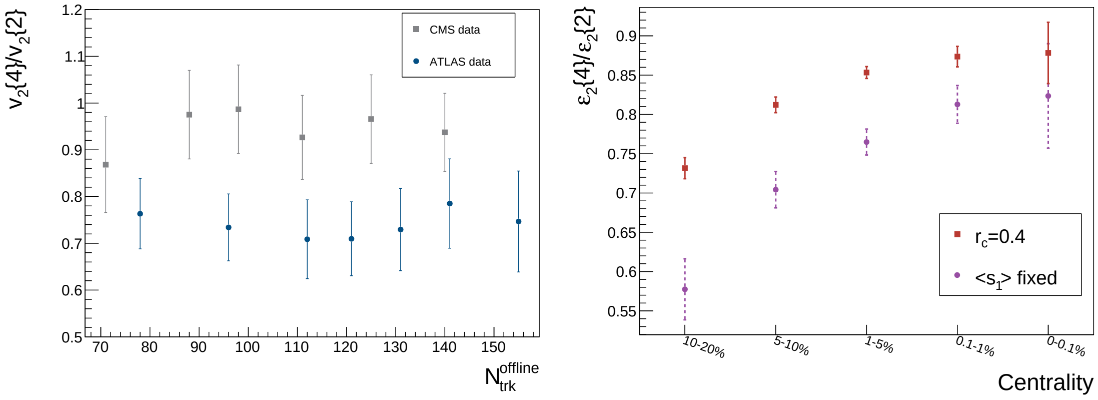

Figure 4.13: Left: Fluctuations of v_2 characterized by the ratio $v_2\{2\}/v_2\{4\}$ as a function of the number of reconstructed charged particles using the CMS definition. Data measured by ATLAS [160] and CMS [170]. Right: Fluctuations of ε_2 within our wounded hot spot approach.

and may lead to confusion for the reader. This is so because $\mathrm{N}_{\mathrm{trk}}^{\mathrm{offline}}$ and N_{ch} are not equivalent. Nevertheless, two interesting messages can be extracted from Fig. 4.13. First, the outcome of our calculation is in quantitative agreement with the experimental data and nicely interpolates between the ATLAS and CMS results indicating that our model although unsophisticated is realistic. The fluctuations of ε_2 turn out to be very sensitive to the presence or absence of correlations, even more than its absolute value as one can observe by comparing Figs. 4.13 and 4.11. Therefore we conclude, as already suggested by [210], that the description of $v_2\{2\}/v_2\{4\}$ may set strong constraints in any initial state model.

4.3.4 Energy scan in minimum bias

In addition to the effect of correlations on the properties of the initial state we have explored their energy dependence from ISR to the LHC. In Fig. 4.14 we represent the average values of ε_2 and ε_3 as a function of their standard deviation for 3 different collision energies namely $\sqrt{s} = 52.6, 7000, 13000$ GeV. All the curves refer to the $r_c = 0.4$ case and all events have been included. It should be noted that the energy dependence of the parameters of the model (see Table 4.1) comes from the requirement of reproducing the total p+p cross section, being this a quite soft condition. Endowing our model with a more rigorous and precise energy dependence is left for future work. The main goal of Fig. 4.14 is to show that there are no significant differences in the values of $\langle\varepsilon_2\rangle$ and $\langle\varepsilon_3\rangle$ for different collision energies. The fact that the spatial eccentricities do not drastically deviate with increasing energy was also observed in [211] where a different parametrization of the initial state was used. Further, the purpose of representing the average values as a function of the standard deviation is to emphasize the width of the probability distributions that we obtain as we have seen in the previous sections (see Figs. 4.7-4.10).

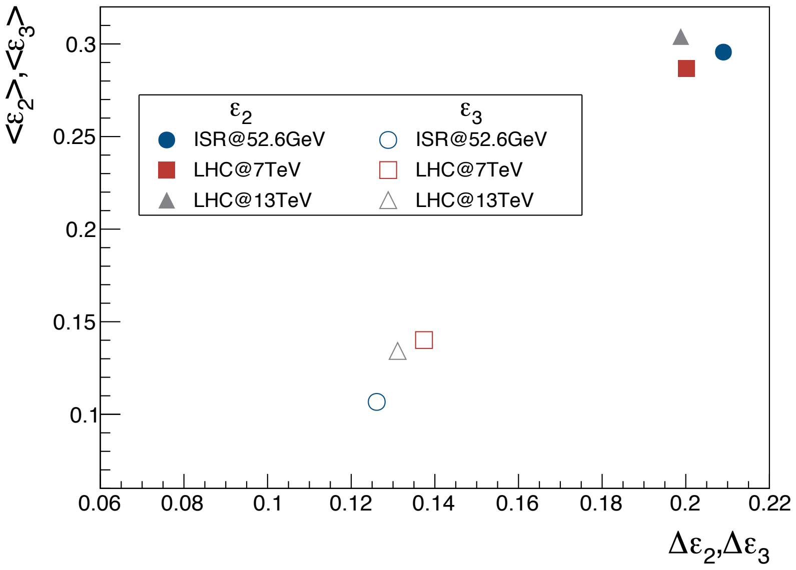

Figure 4.14: Average values of ε_2 (filled markers) and ε_3 (empty markers) for $\sqrt{s} = 52.6$ GeV (blue filled/empty circles), 7 TeV (red filled/empty squares) and 13 TeV (grey filled/empty triangles) as a function of their standard deviation. These energies correspond to ISR and Runs I-II of the LHC, respectively.

	Correlated				**$\langle s_1 \rangle$ fixed**
	R_{hs} [fm]	R [fm]	r_c [fm]	R_p [fm]	R [fm]
$\mathbf{N_{hs}=2}$	0.51	1.04	0.35	1.31	1.13
$\mathbf{N_{hs}=4}$	0.21	0.55	0.32	1.2	0.64

Table 4.3: Default values of the parameters characterizing the hot spots distribution Eq. (4.2) and their probability to interact Eq. (4.3) for different number of hot spots both in the correlated and "$\langle s_1 \rangle$ fixed" cases.

	Correlated				
	$\overline{n}_1$	κ_1	κ_2	$\overline{n}_2$	α
$\mathbf{N_{hs}=2}$	26.22	1.21	4.64	1.79	0.45
$\mathbf{N_{hs}=4}$	25.68	1.46	4.54	1.82	0.34

	$\langle \mathbf{s_1} \rangle$ **fixed**				
	$\overline{n}_1$	κ_1	κ_2	$\overline{n}_2$	α
$\mathbf{N_{hs}=2}$	24.66	1.05	4.55	2.05	0.49
$\mathbf{N_{hs}=4}$	23.04	1.15	4.37	2.14	0.4

Table 4.4: Default values of the parameters of the double Gamma distribution that characterizes the fluctuating amount of entropy each wounded hot spot deposits, s_0, for different number of hot spots both in the correlated (top) and "$\langle s_1 \rangle$ fixed" cases (bottom).

4.3.5 N_{hs}-dependence

All along this Section we have considered that the proton is constituted by 3 gluonic hot spots. This is the most natural scenario when a direct correspondance between the Fock space of valence partons and the hot spots is assumed. However, this relation is arguable as, while being extensively used as a phenomenological tool, the ultimate dynamical origin of the hot spots remains as an open debate (see Sec. 2.3.1). Therefore, it is opportune to check the reliability of our results after variations of this parameter, N_{hs}. Note that the possibility of having a different number of hot spots may account for other elements such as the large-x sea quarks.

We focus our discussion on the results for $\varepsilon_{2(3)}$ after considering the two more straightforward extensions of our model: $N_{hs}=2$ and $N_{hs}=4$. In order to make a fair comparison between the three different scenarios, $N_{hs}=(2,3,4)$, we choose representative values of the parameters $\{R_{hs}, R\}$ that fulfill two constraints. As in the previous sections, the experimental value of the total p+p cross section is reproduced. Further, the proton radius defined as $R_p=\sqrt{N_{hs}}\sqrt{\langle s_1^2\rangle + R_{hs}^2}$, where $\langle s_1 \rangle$ is the r.m.s of the spatial probability distribution given by Eq. (4.2), should not depend on the number of hot spots that the proton contains so we fix it to be the same in all cases. For simplicity, we restrict the analysis to the correlated and $\langle s_1 \rangle$ fixed cases at $\sqrt{s}=13$ TeV. The values of the parameters are given in Tables 4.3 and 4.4. From a technical point of view, the extension of the code to $N_{hs}>3$ brings with it a change in the sampling procedure of the spatial distribution of hot spots $D(\{s_i\})$. This is so because the built-in ROOT class for random number generation extends up to 3 dimensional functions. To overcome this difficulty we have developed an independent algorithm based on the well-known rejection sampling method. The algorithm has been tested against the ROOT routine by generating random numbers according to a 3-dimensional Gaussian distribution. As shown in Fig. 4.15, the accuracy of the method is satisfactorily compatible with the well tested ROOT class. Nevertheless, the current implementation of the Monte Carlo event generator is not straightforwardly extendable to an arbitrary number of hot spots in each proton. From a numerical point

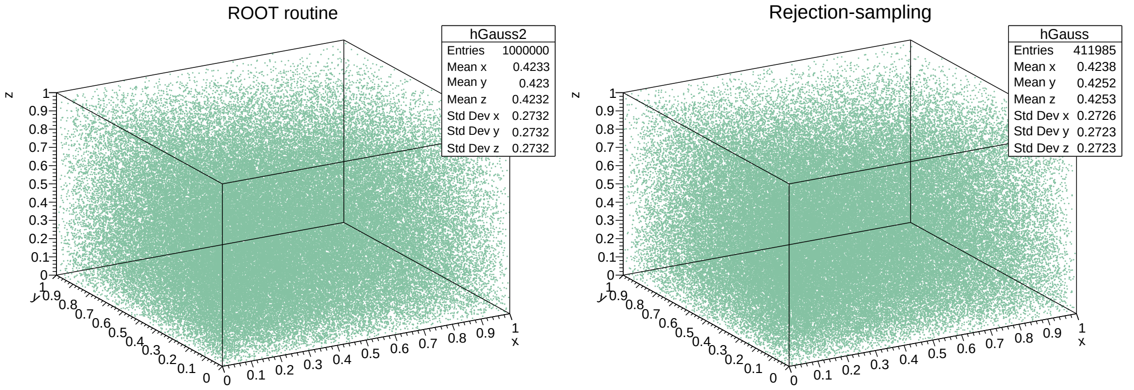

Figure 4.15: Each green dot represents an (x,y,z)-trio distributed according to a 3-D Gaussian distribution. They have been generated with the TRandom class in ROOT (left) and with our own rejection-sampling based algorithm (right). The boxes show the mean values and the standard deviations of (x,y,z).

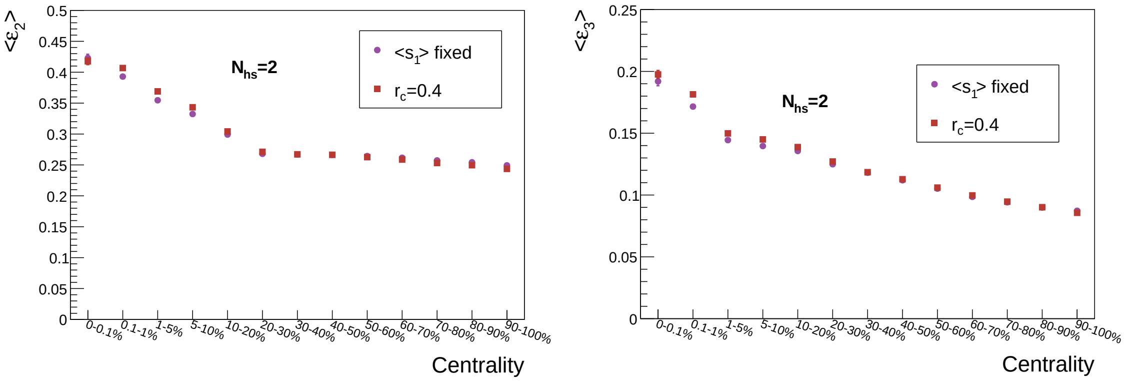

Figure 4.16: Average values of the eccentricity, ε_2, (left) and the triangularity, ε_3, (right) for $N_{hs} = 2$ as a function of the centrality range.

of view, the main bottleneck is the generation of $\mathcal{P}(N_{\rm ch})$ as the number of terms entering the sum increases very rapidly. Up to now, all these possible combinations are computed in a rather brute-force fashion and, therefore, obtaining $\mathcal{P}(N_{\rm ch})$ for a large number of N_{hs} would be doable but computationally expensive. On the analytic side, tuning the model parameters by ensuring that the total cross section is reproduced and the proton radius remain constant makes the full study we are undertaking more cumbersome as meeting both conditions require $3N_{hs}$-dimensional Gaussian integration. Enlarging the flexibility of the code to accomodate an arbitrary number of constituent hot spots is a must for the future. However, even the two simplest extensions considered in the following already show very interesting physical features. The average values of ε_2 and ε_3 against the centrality of the collision for $N_{hs} = 2, 4$ are shown in Fig. 4.16 and 4.17, respectively. Note that we have extended our calculations to a higher centrality bin, $[0-0.1\%]$. First and foremost, the effect of including spatial correlations is invariant under changes in the number of

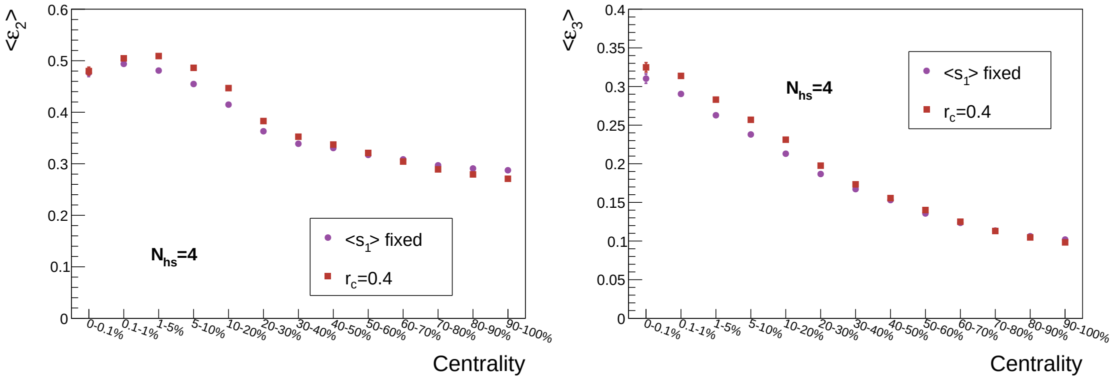

Figure 4.17: Average values of the eccentricity, ε_2, (left) and the triangularity, ε_3, (right) for $N_{hs}=4$ as a function of the centrality range.

hot spots: larger eccentricities and triangularities are expected in the correlated scenario than in the uncorrelated one in ultra-central collisions while they reduce both of them in minimum bias events. Furthermore, we find that both ε_2 and ε_3 smoothly increase when a larger number of hot spots is considered.

4.4 Normalized symmetric cumulants

This Section constitutes the natural extension of our previous studies on the initial state properties of proton-proton collisions in terms of eccentricities by exploring not only their mean but their fluctuations (see Sec. 3.5). There have been attempts to describe the values of the symmetric cumulants, SC(n,m) in p+Pb by computing them in terms of eccentricities, i.e. replacing $v_{n(m)}$ by $\varepsilon_{n(m)}$ in Eqs. (3.38-3.39), within wounded quark models. These studies lead to the correct negative sign of SC(2,3) at high multiplicities but the magnitude is off [212, 193]. Up to today we are unaware of any theoretical prediction for the values of SC(n,m) in p+p interactions at LHC energies, although results for RHIC energies were presented in [212]. All along this Section we compute the normalized version of the symmetric cumulants at the eccentricity level as given by:

$$\mathrm{NSC(n,m)} \equiv \frac{\langle \varepsilon_n^2 \varepsilon_m^2 \rangle - \langle \varepsilon_n^2 \rangle \langle \varepsilon_m^2 \rangle}{\langle \varepsilon_n^2 \rangle \langle \varepsilon_m^2 \rangle} \tag{4.20}$$

at $\sqrt{s} = 13$ TeV, the energy at which they have been measured. The computation of higher order moments of the eccentricities distributions requires more statistics. More concretely, within our Monte Carlo approach, an order of magnitude more of events was needed relative to the eccentricities calculation to achieve enough statistical significance (4.5 million vs 500k). For simplicity, we have not considered the $r_c = 0, nc$ case. In the following plots the uncorrelated results will be exhibited as a band bounded by the $r_c = 0$ and $\langle s_1 \rangle$ fixed cases to display the different possibilities considered. Again, the proton is considered to be formed by 3 hotspots by default and the sensitivity of the results to this parameter is explored in a dedicated section.

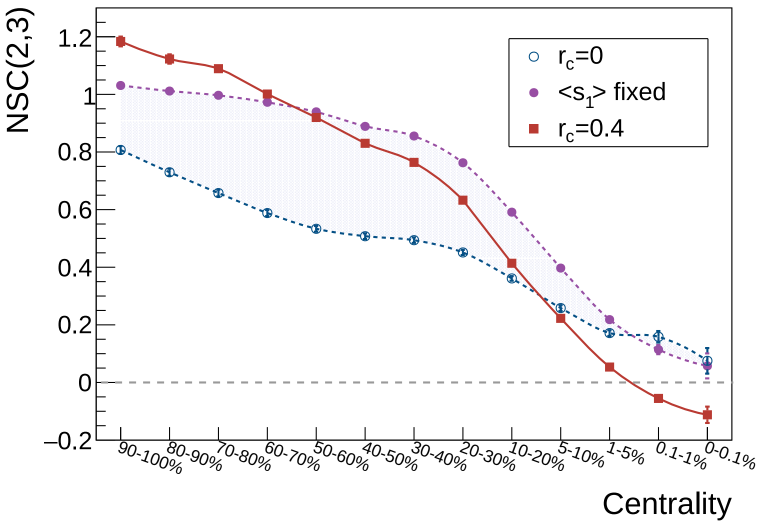

Figure 4.18: Average value of NSC(2,3) as a function of the centrality range for $r_c = 0$ (blue short-dashed line connecting open blue circles), $\langle s_1 \rangle$ fixed (purple short-dashed line connecting filled purple circles) and $r_c = 0.4$ fm (red solid line connecting filled red squares). The error bars represent statistical uncertainties while the light violet band indicates the theoretical uncertainty associated to the choice of parameters that define the uncorrelated scenario.

4.4.1 Centrality dependence

The most important result of this Section is shown in Fig. 4.18 where we represent the event-averaged value of NSC(2,3) as a function of centrality. A common feature in the three correlation scenarios is the fact that NSC(2,3) decreases from peripheral to central collisions as suggested by data. Focusing on the effect of the short-range repulsive correlations we observe how they enlarge the positive correlation of ε_2 and ε_3 in the peripheral regime. However, their repercussion in the very central collisions is precisely the opposite. Finally, the most striking effect of the spatial correlations is observed in the ultra-central bins $[0 - 0.1\%]$ and $[0.1 - 1\%]$: only in the $r_c = 0.4$ case there exists an anti-correlation of ε_2 and ε_3 as data dictates. Then, we conclude that the experimental evidence of NSC(2,3)< 0 may back up the necessity to consider correlated proton constituents. Note that a precise theory-to-data comparison, is not possible due to the aforementioned multiplicity selection. Especially cumbersome is the comparison in the low multiplicity regime where non-flow dijet contributions, totally absent in our initial state coordinate space approach, dominate the measured values of NSC(n,m). However, our main purpose is to lay out, for the first time in the literature, a particular mechanism i.e. the presence of spatial repulsive correlations inside the proton, that builds up a negative sign of NSC(2,3) in the highest centrality bin at the geometric level.

In the case of NSC(2,4), see Fig. 4.19, the role of the repulsive correlations is qualitatively the same as in the NSC(2,3) calculation: in peripheral collisions the value of

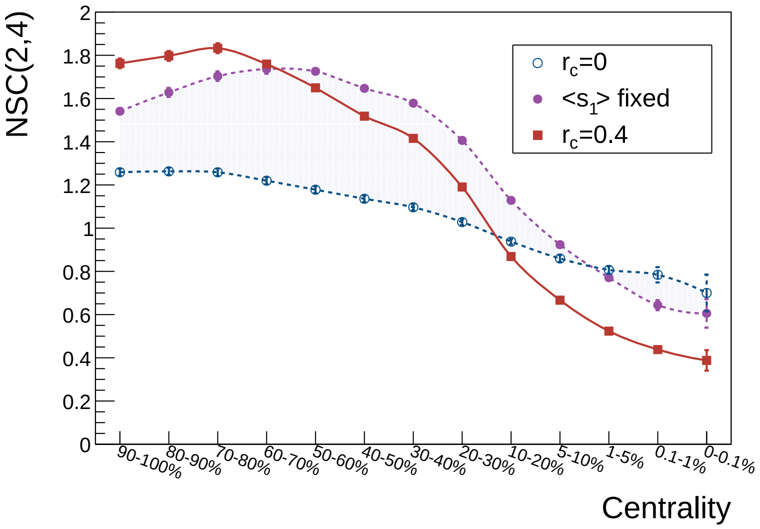

Figure 4.19: Average value of NSC(2,4) as a function of the centrality range for $r_c = 0$ (blue short-dashed line connecting blue open circles), $\langle s_1 \rangle$ fixed (purple short-dashed line connecting filled purple circles) and $r_c = 0.4$ fm (red solid line connecting filled red squares). The error bars represent statistical uncertainties while the light violet band indicates the theoretical uncertainty associated to the choice of parameters that define the uncorrelated scenario.

NSC(2,4) is larger in the $r_c = 0.4$ case than in the uncorrelated scenarios and the situation gets reversed at barely the same centrality bin. As well, we find the absolute value of NSC(2,4) to be larger than NSC(2,3) in all the centrality bins as it is the case in the data. We would also like to remark that in our approach the symmetric cumulants are almost flat in the mid-to-peripheral interactions but thanks to a dissection of the very central bins we see a clear centrality dependence. This is consistent with our previous calculations of the average values of the spatial eccentricity moments (see Figs. 4.11-4.17).

A geometric and intuitive interpretation of the fact that only in the correlated case NSC(2,3)< 0 in the $[0-1\%]$ centrality bin is given in the following section. It should be noted that, for this purpose, we have merged the two highest centrality bins, $[0-0.1\%]$ and $[0.1-1\%]$, into a single one in order to improve the statistics.

4.4.2 Role of the interaction topology

In order to capture the effect of the spatial correlations we characterize each proton-proton interaction by its number of wounded hot spots and the number of collisions $(N_w, N_{\rm coll})$, the two basic quantities of any Monte Carlo Glauber calculation. We dub each $(N_w, N_{\rm coll})$-configuration as *interaction topology*.

We begin our analysis by computing the average number of collisions as a function of the number of wounded hot spots for the three different scenarios introduced above.

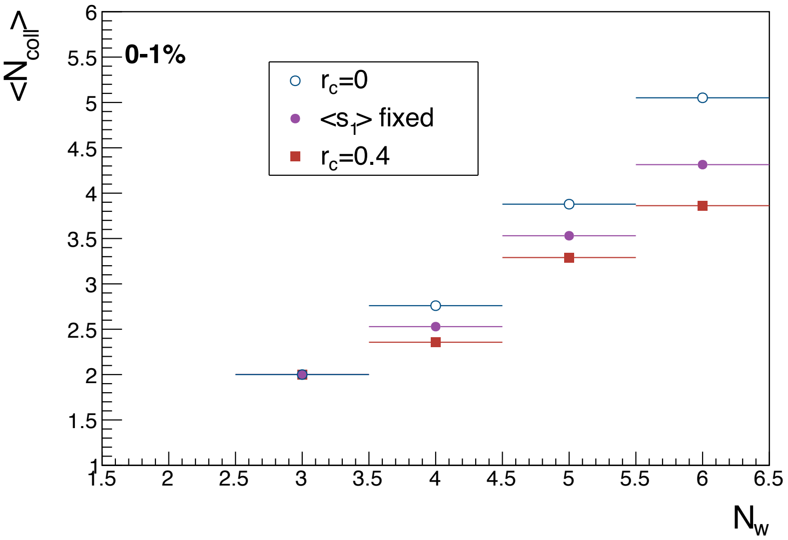

Figure 4.20: Average number of collisions as a function of the number of wounded hot spots for $r_c = 0$ (open blue circles), $\langle s_1 \rangle$ fixed (filled purple circles) and $r_c = 0.4$ fm (filled red squares).

The results are shown in Fig. 4.20. First of all, as we describe the entropy deposition in an incoherent way i.e. on average the more wounded hot spots the more entropy is deposited, the configurations in which only two hot spots collide cannot create enough entropy to be part of the $[0-1\%]$ centrality bin. Then, the minimum number of wounded hot spots is three and, in this case, $\langle N_{\rm coll} \rangle = 2$ in all the correlation scenarios as it is the only existing configuration. However, for $N_w > 3$ the average number of collisions starts to differ between the three different cases. We observe that $\langle N_{\rm coll} \rangle$ is systematically reduced when including repulsive correlations with respect to the uncorrelated cases. This effect has a very straightforward interpretation: enlarging the mean transverse distance between the hot spots reduces the probability of having interaction topologies with a high number of collisions. In other words, the repulsive correlations spread the hot spots in the transverse plane and as a consequence enhance the probability of the hot spots to collide by pairs over the configurations in which all hot spots in the projectile interact with all the others in the target, as it is schematically represented in Fig. 4.21.

To connect this fact with the total value of NSC(2,3) we would like to understand the individual contributions from the different interaction topologies. For this purpose we define a weighted version of NSC(n,m) denoted $\mathrm{NSC_w}$(n,m) as follows

$$\mathrm{NSC_w(n,m)} \equiv \mathcal{P}(N_w) \cdot \mathcal{P}(N_w|N_{\rm coll}) \cdot \mathrm{NSC(n,m)}\Big|_{N_w,N_{\rm coll}} \tag{4.21}$$

where

- $\mathcal{P}(N_w)$ is the probability of having a certain number of wounded hot spots.

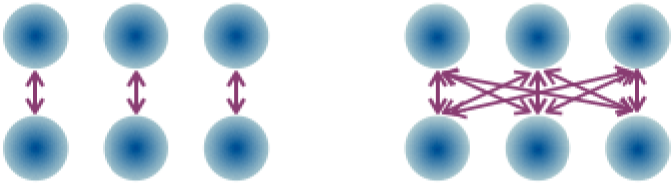

Figure 4.21: Sketch representing the interaction topologies preferred in the correlated case (left) and in the uncorrelated one (right). The purple arrows represent the collisions between the hot spots.

- For a given N_w, $\mathcal{P}(N_w|N_{\rm coll})$ represents the probability of having a certain number of collisions between the hot spots.
- $\mathrm{NSC(n,m)}\Big|_{N_w,N_{\rm coll}}$ is the value of NSC(n,m) for each interaction topology.

The error of NSC$_{\rm w}$(n,m) is computed by adding the statistical uncertainties of each term in Eq. (4.21) in quadrature. Essentially, by summing NSC$_{\rm w}$(n,m) over all the possible $(N_w, N_{\rm coll})$-configurations one recovers NSC(n,m). This new quantity allows us to decompose the value of NSC(2,3) and investigate the contribution of each interaction topology separately. From now on, to facilitate the discussion, we only show the comparison between $\langle s_1\rangle$ fixed and $r_c = 0.4$ scenarios. We have checked that the same conclusions as in the $\langle s_1\rangle$ fixed case hold for $r_c = 0$.

In Fig. 4.22 we show a particular example of the output of our calculation for NSC$_{\rm w}$(2,3) by selecting the events with $N_w = 6$. Two important results can be extracted from this figure. First, as already suggested by Fig. 4.20, configurations with a large number of collisions, e.g. $N_{\rm coll} > 6$, only occur in the uncorrelated case where the three hot spots are closer to each other or, equivalently, clustered. Second, and more important, the value of NSC$_{\rm w}$(2,3) shows a clear dependence on $N_{\rm coll}$: configurations with a smaller number of collisions reduce the value of NSC(2,3) and, eventually, contribute negatively. Then, in our picture, the inclusion of spatial correlations inside the proton modifies the weight of each interaction topology in such a way that these configurations are enhanced. This feature provides a natural explanation for the different sign of NSC(2,3) in the uncorrelated and correlated scenarios. Far from being a casual coincidence or an artifact this effect is observed for any number of wounded hot spots as it is depicted in Fig. 4.23. In the left pannel we show the event-averaged value of NSC$_{\rm w}$(2,3) with respect to the number of collisions for $N_w = 3$ to $N_w = 6$ for the $r_c = 0.4$ case. Once again, the configurations that contribute more to the total value of NSC(2,3) are the ones with a large number of wounded hot spots that interact a small amount of times. In opposition, as displayed in the right pannel, the interaction topologies that have associated a negative NSC(2,3) are extremely suppressed in the uncorrelated scenario where the configuration with the biggest weight and precisely positive value of NSC(2,3) is $(N_w = 4, N_{\rm coll} = 3)$.

Then, by computing NSC$_{\rm w}$(2,3) for the different interaction topologies we find that the origin of the negative sign of NSC(2,3) in the $r_c = 0.4$ scenario is due to the decisive role

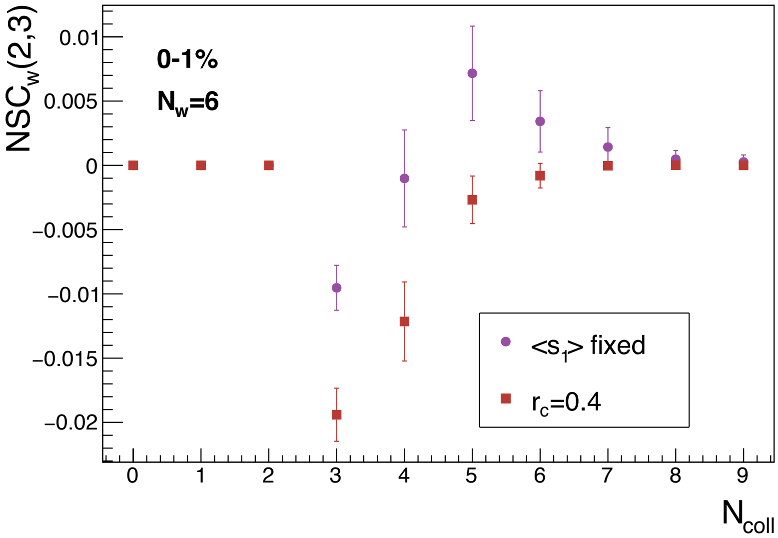

Figure 4.22: Average value of NSC$_\mathrm{w}$(2,3) as a function of the number of collisions after selecting the events with $N_\mathrm{w} = 6$ for $\langle s_1 \rangle$ fixed (filled purple circles) and $r_c = 0.4$ fm (filled red squares).

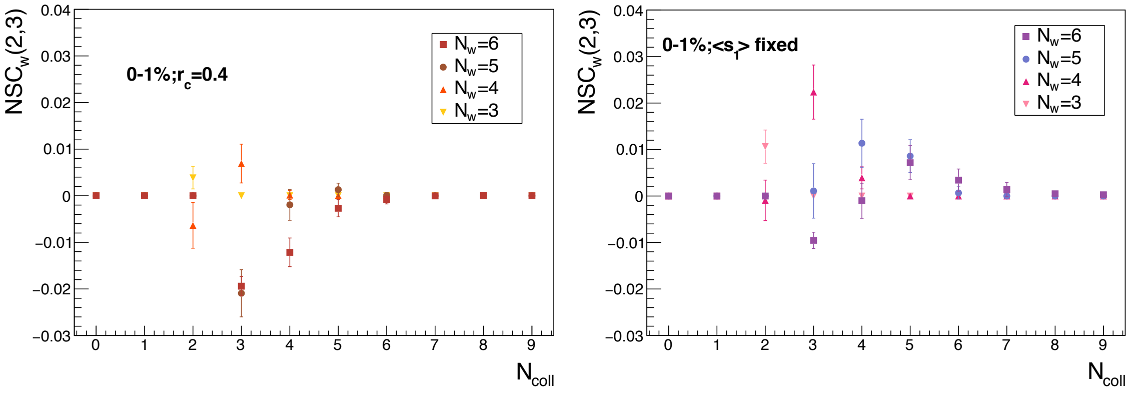

Figure 4.23: Average value of NSC$_\mathrm{w}$(2,3) as a function of the number of collisions for different number of wounded hot spots. Left: $r_c = 0.4$ fm case. Right: $\langle s_1 \rangle$ fixed case.

of correlations in modifying the weights of the diverse configurations in the Monte Carlo Glauber simulations.

4.4.3 Scan of the parameter space

To continue, we check the sensitivity of the obtained results on the values of the model parameters. Thus, we focus on the correlated scenario and study the dependence of NSC(2,3) on the radius of the hot spot and the repulsive core distance in the $[0-1\%]$ centrality bin. The value of the repulsive distance, r_c, apart from being different from zero [59], is essentially unconstrained. As it could be argued that $r_c=0.4$ fm is a large repulsive distance that may be unrealistic we explore the results of our model for $r_c=0.25$ fm. In the case of R_{hs}, we choose 4 different values in our scan $\{0.15, 0.25, 0.32, 0.4\}$ fm. Consequently, the parameters of the Gamma distribution for the entropy deposition (see Eq. (4.8)) are extracted in all the cases by fitting the experimental charged-particle multiplicity distributions $\mathcal{P}(N_{\rm ch})$. The other two parameters of our model, namely R and ρ_{hs}, remain fixed to their default values given in Table 4.1. Except in the chosen values for R_{hs} and r_c appearing in Table 4.1, i.e. $R_{hs}=0.32$ fm and $r_c=0.4$ fm, the requirement that our model reproduces the p+p total cross section and ρ is not fulfilled. Removing these phenomenological constraints allows to pinpoint the effect of just varying the radius of the hot spot or the correlation distance in our results.

In the left pannel of Fig. 4.24 we represent the event-averaged value of NSC(2,3) as a function of R_{hs} for the two different values of the correlation distance considered. First, we observe that by reducing the value of r_c for a given value of R_{hs} we get closer to the uncorrelated case and thus the value of NSC(2,3) is enlarged and pushed to the positive regime, as expected. However, this statement is not universal as it breaks down when $R_{hs} \lesssim 0.22$ fm. In this scenario of very small values of the radius of the hot spot, ε_2 and ε_3 are positively correlated for both values of the repulsive core distance and the value of NSC(2,3) is larger in the $r_c=0.4$ case. This result indicates that NSC(2,3) is not sensitive to R_{hs} and r_c independently but to the interplay of both scales. In other words, NSC(2,3) depends on a generic function of the radius of the hot spot and the repulsive core distance $f(R_{hs}, r_c)$.

As a first and simple guess to the functional form of $f(R_{hs}, r_c)$ we choose it to be the ratio of the two scales involved i.e. $f(R_{hs}, r_c)=R_{hs}/r_c$. This ratio has a transparent interpretation by characterizing the degree of repulsion: if $R_{hs}/r_c \gg 1$ the shape of the proton resembles the uncorrelated scenario where the hot spots can largely overlap in transverse space. The results of NSC(2,3) for different values of R_{hs}/r_c are displayed in the right pannel of Fig. 4.24. We can distinguish three regimes in this plot. On the one hand, when $R_{hs}/r_c \geq 1$ the geometric picture of the proton approaches the uncorrelated scenario and the value of NSC(2,3) increases monotonically starting to be positive for $R_{hs}/r_c \gtrsim 1.3$. Moreover, when $R_{hs}/r_c=1$ the value of NSC(2,3) is identical within error bars for both correlation scenarios $r_c=0.25$ and $r_c=0.4$ supporting the idea that NSC(2,3) depends on $f(R_{hs}, r_c)=R_{hs}/r_c$. However, on the second regime characterized by $0.6 \lesssim R_{hs}/r_c \lesssim 1$, we find an abrupt change of the value of NSC(2,3) when slightly increasing the ratio R_{hs}/r_c from 0.6 to 0.63. This suggests that a residual dependence of NSC(2,3) on the

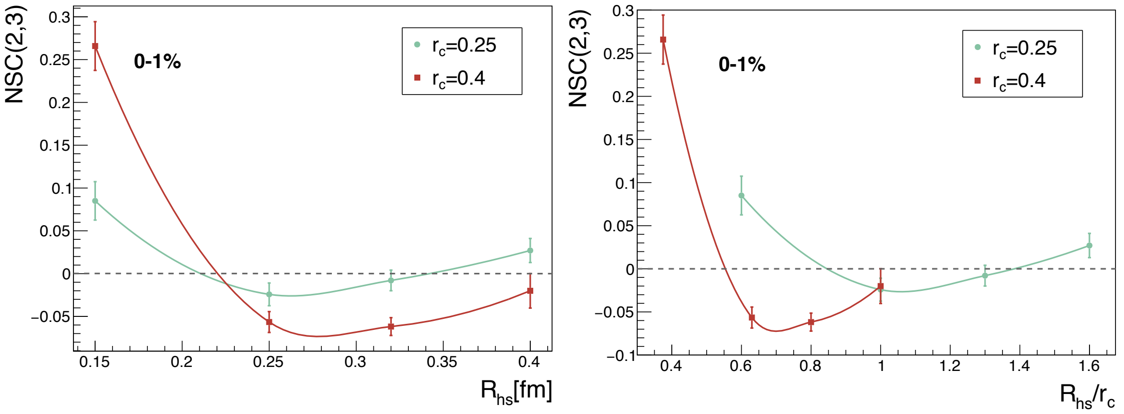

Figure 4.24: Average value of NSC(2,3) for two different values of the repulsive distance: $r_c = 0.25$ fm (filled green circles) and $r_c = 0.4$ fm (filled red squares) as a function of the radius of the hot spot (left) and the ratio R_{hs}/r_c (right).

other scale of the problem, R, may exist. Finally, configurations in which the hot spots are much smaller than the repulsive core distance between them i.e. $R_{hs}/r_c \lesssim 0.6$ result into a positive correlation between ε_2 and ε_3. Then, our study favors values of $0.6 \lesssim R_{hs}/r_c \lesssim 1.3$ in order to be compatible with the experimental observation of NSC(2,3)<0 in the highest centrality bin. Unfortunately, this interval is large enough to be compatible with a picture of the proton in which the hot spots transverse separation is larger than in the uncorrelated case but still they can overlap ($R_{hs}/r_c \sim 1.3$) and with a much more dilute description in which the probability of two hot spots to overlap is highly supressed ($R_{hs}/r_c \sim 0.6$).

4.4.4 Sensitivity of NSC(2,3) to N_{hs}

As in the case of the eccentricities, we check the robustness of our results after variations in the number of constituent hot spots. We start by exploring the dependence of the event-averaged value of NSC(2,3) on the number of hot spots in the correlated scenarios as displayed in Fig. 4.25 as a function of centrality. The differences between the three cases start to appear in mid-to-ultra central collisions. There exists a clear trend towards smaller values of NSC(2,3) when a bigger number of hot spots is considered. Specifically, the negative sign of NSC(2,3) in the high centrality bins is not achieved when $N_{hs} = 2$ even with correlations. Thus, we conclude that with the selected parameters the minimum number of hot spots to describe the onset of the anti-correlation between ε_2 and ε_3 is $N_{hs} = 3$. Note that this threshold number of hot spots coincides with the one needed to describe the onset of the hollowness effect (see Chapter 2). Moreover, the inclusion of an additional hot spot i.e. $N_{hs} = 4$ helps to make NSC(2,3) even more negative in the highest centrality bins although the effect is small when compared to the drastic impact of changing from $N_{hs} = 2$ to $N_{hs} = 3$.

In Fig. 4.26 the comparison between the correlated and $\langle s_1 \rangle$ fixed scenarios for $N_{hs} = 2$ (left) and $N_{hs} = 4$ (right) is displayed. First and foremost, the effect of including spatial correlations is invariant under changes in the number of hot spots: in peripheral colli-

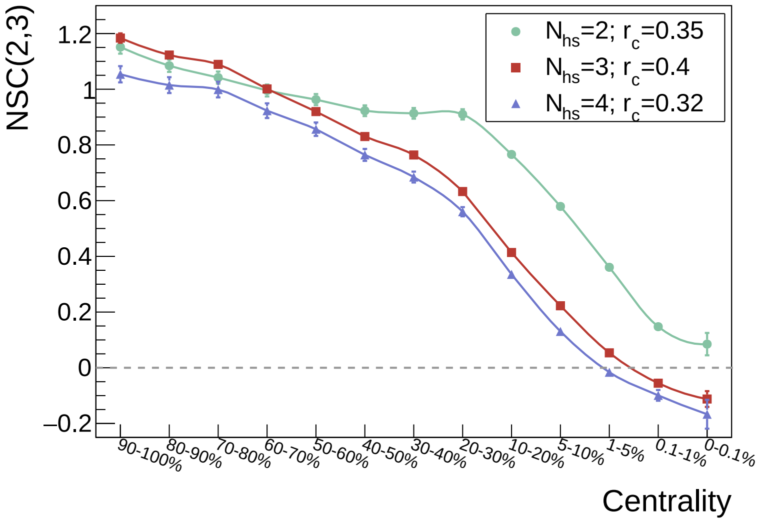

Figure 4.25: Average value of NSC(2,3) as a function of the centrality range in the correlated scenario for $N_{hs}=2$ (filled green circles), $N_{hs}=3$ (filled red squares) and $N_{hs}=4$ (filled violet triangles). The error bars represent statistical uncertainties.

sions they enlarge the positive correlation of ε_2 and ε_3 while favoring a negative sign of NSC(2,3) with respect to the uncorrelated scenario. Albeit the correlated curve is always below the uncorrelated scenario in the highest centrality bins an important comment is in order: NSC(2,3) is compatible with negative values, within statistical uncertainty, in the $[0-0.1\%]$ bin for the uncorrelated case. This fact reinforces the idea remarked in the previous section: the interplay of the different scales $\{R_{hs}, r_c, N_{hs}\}$ is decisive in the sign of NSC(2,3) within our framework. For $N_{hs}=4$ the weight of the configurations with a large number of wounded hot spots and a small number of collisions is large enough so that the spatial correlations are not essential to obtain a negative NSC(2,3) in the $[0-0.1\%]$ bin. However, as these configurations are enhanced in the correlated scenario, the anti-correlation of ε_2 and ε_3 is stronger than in the uncorrelated case just as in the case of $N_{hs}=3$. To sum up, although the negative sign of NSC(2,3) in the highest centrality bins is not a unique feature of the correlated scenario but relies on the interplay of the different scales, the inclusion of repulsive correlations provides a mechanism to reduce its value in the highest centrality bins.

4.4.5 Influence of N_w/N_{coll} on NSC(2,3)

All along the previous sections we have reached the firm conclusion that NSC(2,3) is extremely sensitive to the initial state fluctuations and can help to discriminate between different parameterizations of the proton geometry. However, a natural question remains unanswered: is there a physical parameter controlling the sign of NSC(2,3)? Being our model a multiparametric one the task of pinpointing the critical parameter is far from being straightforward as we have seen in previous attempts by testing the sensitivity to

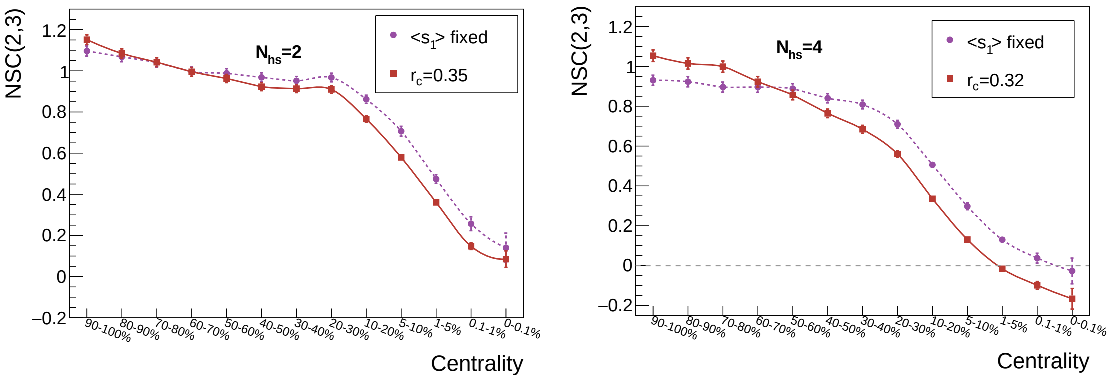

Figure 4.26: Average value of NSC(2,3) as a function of the centrality range for $\langle s_1 \rangle$ fixed (purple short-dashed line connecting filled purple circles) and $r_c = 0.4$ fm (red solid line connecting filled red squares). The error bars represent statistical uncertainties. Left: $N_{hs} = 2$. Right: $N_{hs} = 4$.

the different scales of the problem. The study in terms of the interaction topologies guided our choice to N_w/N_{coll} as a potential candidate.

Thus, we study the dependence of NSC(2,3) on this ratio for the highest centrality bins with and without correlations in all cases where R_p is identical. The results are shown on the left pannel of Fig. 4.27. Remarkably, the value of the cumulant is the same in all the different scenarios for a given value of N_w/N_{coll}. Nor the presence/absence of correlations neither the number of constituent hot spots modify the value of NSC(2,3) for a fixed N_w/N_{coll} ratio. This feature backs up the relevant role of this parameter in the values of NSC(2,3). To facilitate the discussion on the sign of NSC(2,3) we identify two separated regimes. For $N_w/N_{\text{coll}} < 1$, the results are dominated by statistical uncertainties as it is rather unusual that the number of collisions surpasses the number of participants. At first sight, the small error bars and neat negative sign of the point sitting at $N_w/N_{\text{coll}} = 0.7$ for the $N_{hs} = 4$; $\langle s_1 \rangle$ fixed case is, perhaps, surprising. A deeper look into this value reveals that it corresponds to the configuration ($N_w = 7$,$N_{\text{coll}} = 10$) that occurs only three times in the Monte Carlo data set used to obtain this plot. The reduced statistical uncertainty in such a fluctuating quantity as NSC(2,3) indicates that this point may be an outlier. More statistics is needed to clarify this aspect. In the opposite regime of $N_w/N_{\text{coll}} > 1$, a general trend is observed. The sign of NSC(2,3) tends to negative values when enlarging the value of N_w/N_{coll} reaching its maximum for $N_w/N_{\text{coll}} = 2$ i.e. when the wounded hot spots interact by pairs. Therefore, we identify these configurations to be responsible of the negative sign of NSC(2,3) in the highest centrality bins. Then, the main distinction between the different cases considered in this work is the probability of these configurations to happen. For example, when $N_{hs} = 4$ the weight of configurations with large N_w/N_{coll} ratio is enhanced with respect to the $N_{hs} = 3$ case and then permit a negative value of NSC(2,3) even within the uncorrelated scenario.

To check to what extent the previous conclusions depend on the size of the proton we repeat the computation of NSC(2,3) as a function of N_w/N_{coll} for $N_{hs} = 3$ and consider three different situations. In addition to the default $r_c = 0.4$ fm and $\langle s_1 \rangle$ fixed cases, we

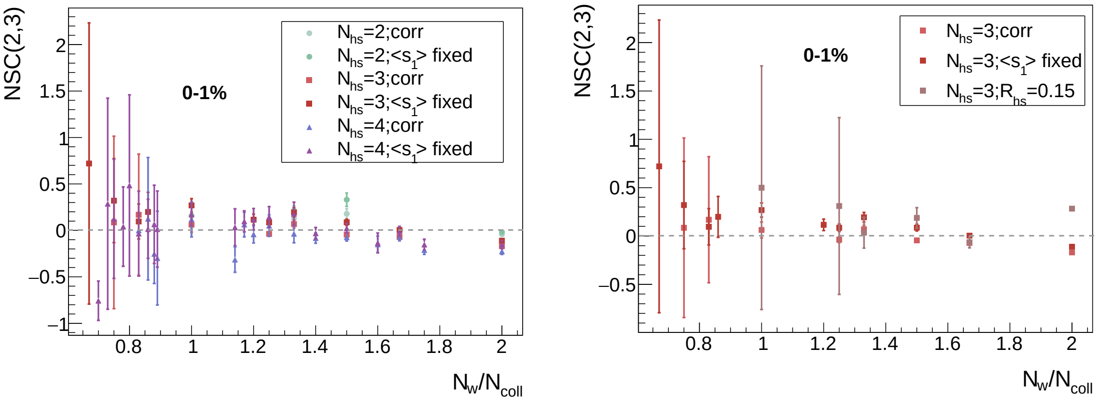

Figure 4.27: Average value of NSC(2,3) as a function of $N_w/N_{\rm coll}$. Left: $N_{hs} = (2, 3, 4)$ for the correlated and uncorrelated scenarios. Right: $N_{hs} = 3$ for the $\langle s_1 \rangle$ fixed case and the correlated one with two different values of $R_{hs} = (0.15, 0.32)$.

reduce R_{hs} down to 0.15 fm by keeping the same correlation distance as it was done in Fig. 4.24. From the right side of Fig. 4.27 one sees how the value of NSC(2,3) does not coincide in the three scenarios considered as was the case on the left side of the plot. Therefore, the *universality* of NSC(2,3) as a function of $N_w/N_{\rm coll}$ requires the global size of the proton to be identical in the different configurations to be compared.

Although the previous results encourage to identify $N_w/N_{\rm coll}$ as the physical parameter controlling the sign of NSC(2,3), this is not the end of the story. As an example, an interesting question that requires further investigation is whether these arguments hold, and if so, how do they relate to the p+p scenario, in an asymmetric collision such as p+Pb where, if the proton substructure is not considered, the ratio $N_w/N_{\rm coll}$ is always smaller than one.

Finally;
Franz Ferdinand.

Conclusions and future prospects

Up to now it has not been possible to firmly establish a full, self-consistent, multi-dimensional map of the distributions of partons inside the proton in space, momentum and spin. The biggest stumbling-block to accessing this information from a theoretical perspective is its non-perturbative nature within QCD. Beyond its crucial role to sharpen our knowledge about the building blocks of matter, the ample amount of highly-precise data being released by collider experiments such as RHIC or the LHC calls for a fine detailed description, at the subnucleonic level, of the proton structure.

In this thesis we address the characterization of the transverse distribution of proton constituents in coordinate space and their fluctuations. We take as starting point the data sample on the elastic differential cross section measured by ISR and LHC experiments. A successful description of bulk features of the data yields a composite picture of the proton in terms of gluonic hot spots that are spatially correlated and whose radius swells with increasing energy. Next, driven by the suggestive signals of quark-gluon plasma formation in p+p collisions at the LHC, we develop a Monte Carlo Glauber event generator to test the sensitivity of generic properties of the geometry of the collision to the presence of spatial correlations. The ensemble of results achieved in this thesis are summarized next.

■ Microscopic realization of the hollowness effect in proton-proton interactions at the LHC.

We propose that the explanation to the rather counterintuitive hollowness effect, whereby proton peripheral collisions are more effective producing new particles than central ones at high energies, lies in the interplay between the different internal scales of the proton: proton radius, hot spot radius and transverse correlation length. The relative enhancement of the destructive interference terms in the multiple scattering series, known as *shadowing* corrections, induced by non-trivial probability distributions for the hot spots transverse positions and the swelling hot spots radius with increasing energy yield the observed depletion of the inelasticity density in central collisions. Three main conclusions can be drawn from this project:

- The hollowness effect, to be confirmed by the $\sqrt{s} = 13$ TeV data, has a quantum mechanical origin. Therefore, a coherent description of the scattering is mandatory and our choice of the Glauber model is well justified.
- Spatial correlations between the constituents of the proton are of most importance in the description of the hollowness effect.

- Within our picture the emergence of the hollowness effect can not be described when the number of hot spots is smaller than 3.

■ Imprints of spatially correlated constituents on the initial geometry of proton-proton interactions within a Monte Carlo Glauber model.

We present a quantitative analysis of the effect of non-trivial spatial correlations between constituents of the proton on the features of the initial state geometry in proton-proton collisions at high energies in terms of spatial eccentricities and symmetric cumulants. This part of the thesis holds the following results:

- We find that both the eccentricity and the triangularity are affected by the inclusion of short-range repulsive correlations. In particular, the correlated scenario yields larger values of $\varepsilon_{2(3)}$ in ultra-central collisions while reducing them in minimum bias.

- The fluctuations of the eccentricity as described by the ratio $\varepsilon_2\{2\}/\varepsilon_2\{4\}$ show a larger degree of sensitivity to the presence of correlations. The outcome of our calculation reproduces the general trends of the experimental data and is compatible with the measured value.

- We have explored the energy dependence of ε_2 and ε_3, encoded in the growth of the hot spot radius, from ISR to the LHC finding small deviations in their values.

- The bulk of our results for $\varepsilon_{2(3)}$ remain qualitatively the same when varying the number of gluonic hot spots from three to $N_{hs} = (2,4)$. A smooth increasing behavior of $\varepsilon_{2(3)}$ when enlarging N_{hs} is observed.

- We lay out, for the first time in the literature, a particular mechanism that permits an anti-correlation of ε_2 and ε_3 in the highest centrality bins as dictated by data. When modeling the proton as composed by 3 gluonic hot spots, the most common assumption in the literature, we find that the inclusion of spatial correlations is indispensable to reproduce the negative sign of NSC(2,3).

- We confirm the large discriminating power of this observable on any realistic initial state model by scanning the parameter space. This results into stringent constraints on the ratio R_{hs}/r_c within our model.

- The values of NSC(2,3) are sensitive to the variation of the number of hot spots that constitute the proton. Specifically, we find that, within the correlated scenario, adding an extra hot spot to our description reinforces the negative sign of NSC(2,3) while reducing it to $N_{hs} = 2$ pushes NSC(2,3) towards positive values. For $N_{hs} = 4$ we obtain a negative value of NSC(2,3) within the uncorrelated scenario although its absolute value is smaller than in the correlated case.

- We interpret this last result by showing a suggestive universal behavior of NSC(2,3) when evaluated as a function of N_w/N_{coll}. The weight of the configurations with a large value of this ratio in the Monte Carlo simulation has a dramatic influence on the

> sign of NSC(2,3). In the correlated scenario, the probability of having interactions with a large number of wounded hot spots colliding a small amount of times is enhanced. These interaction topologies are responsible in our set up of the negative sign of the symmetric cumulant NSC(2,3).

The results presented in this thesis shall be regarded as the first steps towards a broader objective: the dynamical understanding of the origin of spatial correlations inside the proton and their phenomenological implications. To this concern, there are two major lines of research to be explored in the near future.

On one hand, in order to do a one-to-one comparison of our initial state model with the experimental data on flow harmonic coefficients and symmetric cumulants we will feed a relativistic viscous hydrodynamic simulation with the initial entropy density profiles obtained within our Monte Carlo Glauber model. This project will allow us to check if the effect of the spatial correlations is washed out by the hydrodynamic evolution or, on the contrary, gives a better description of the experimental data. If a linear relationship is assumed between the spatial eccentricity moments and the flow harmonic coefficients, the net effect of spatial correlations is expected to be an enhancement of the elliptic and triangular flow in high-multiplicity proton-proton interactions. Further, we would like to calibrate the sensitivity of our results to the system size by extending our framework to include p+Pb collisions. In such small systems, it is a must to demonstrate that our approach does not break down the assumptions underlying the applicability of hydrodynamics.

On the other hand, for the sake of reinforcing the theoretical bedrocks of our model for the proton, we are working towards its explicit realization within the Color Glass Condensate formalism. The introduction of spatial correlations between a reduce number of sources of color charge, i.e. the hot spots, would imply going beyond the approximation taken in the successful McLerran-Venugopalan model of a Gaussian probability distribution for the color charges, justified by the central limit theorem when the number of them is large enough such as in the nuclear case. Thus, the final goal of this study besides its theoretical content is to gauge the impact of a finite number of correlated colour sources in the initial state momentum correlations.

A successful completion of both projects will open the door for a profound understanding of the intriguing experimental hints of quark-gluon plasma droplets production in small collision systems from an initial and a final state perspective.

Finally;
Franz Ferdinand.

Conclusiones y futuras líneas de investigación

Hasta la fecha no ha sido posible establecer firmemente un mapa completo, auto consistente y multidimensional de las distribuciones de partones en el espacio, momento y espín. El mayor obstáculo para acceder a esta información desde una perspectiva teórica es su naturaleza no perturbativa dentro de QCD. Más allá de su papel crucial en aras de afianzar nuestro conocimiento sobre los constituyentes fundamentales de la materia, el análisis de la amplia cantidad de datos de alta precisión obtenidos en colisionadores de partículas como RHIC o el LHC requiere una descripción fina y detallada, a nivel subnuclear, de la estructura del protón.

Esta tesis aborda la caracterización de la distribución transversa de los constituyentes del protón en espacio de coordenadas y sus fluctuaciones. Tomamos como punto de partida el conjunto de datos sobre la sección eficaz diferencial elástica medido por los experimentos ISR y LHC. Una descripción satisfactoria de las características globales de dichas colisiones da lugar a una imagen compuesta del protón en términos de regiones de alta densidad gluónica que llamamos *hot spots*. Estos hot spots están correlacionados espacialmente y su radio crece al incrementar la energía. A continuación, instigados por las señales que sugieren la formación de plasma de quarks y gluones en colisiones p+p en el LHC, desarrollamos un generador de eventos Monte Carlo Glauber para comprobar la sensibilidad de las propiedades genéricas de la geometría de la colisión a la presencia de correlaciones espaciales. El conjunto de resultados logrados en esta tesis se resumen a continuación.

■ Realización microscópica del efecto hollowness en interacciones protón-protón a energías del LHC.

Proponemos que la explicación del más bien contraintuitivo efecto *hollowness*, por el que las colisiones periféricas entre protones son más destructivas que las centrales a altas energías, radica en la interrelación entre las diferentes escalas internas del protón: su radio, el radio de los hot spots y la longitud de correlación transversa. El mayor peso relativo de los términos de interferencia destructiva en la expansión en serie de la interacción, conocidos también como correcciones de apantallamiento, debido a las distribuciones de probabilidad no triviales para las posiciones transversas de los hot spots y el crecimiento de su radio con el aumento de la energía es capaz de generar la observada reducción de la densidad de inelasticidad en colisiones centrales. Se pueden extraer tres conclusiones principales de este proyecto:

- El efecto hollowness, que ha de ser confirmado por los datos a $\sqrt{s} = 13$ TeV, tiene un origen cuántico. Por lo tanto, es necesaria una descripción coherente de la interacción y nuestra elección del modelo de Glauber está bien justificada.

- Las correlaciones espaciales entre los constituyentes del protón son de suma importancia en la descripción del efecto hollowness.
- Dentro de nuestra modelo, no somos capaces de describir la aparición del efecto hollowness cuando el número de hot spots es menor a 3.

■ Impacto de constituyentes espacialmente correlacionados en la geometría inicial de las interacciones protón-protón dentro de un modelo Monte Carlo Glauber.

Presentamos un análisis cuantitativo del efecto de las correlaciones espaciales no triviales entre los constituyentes del protón en las características de la geometría de estado inicial en colisiones protón-protón a altas energías mediante el cálculo de excentricidades espaciales y cumulantes simétricos. Esta parte de la tesis contiene los siguientes resultados:

- Encontramos que tanto la excentricidad como la triangularidad se ven afectadas por la inclusión de correlaciones de repulsión de corto alcance. En particular, el escenario correlacionado resulta en valores mayores de $\varepsilon_{2(3)}$ en colisiones ultracentrales mientras que los reduce en colisiones de tipo *minimum bias*.
- Las fluctuaciones de la excentricidad caracterizadas a través del cociente $\varepsilon_2\{2\}/\varepsilon_2\{4\}$ muestran un mayor grado de sensibilidad a la presencia de correlaciones. El resultado de nuestro cálculo reproduce las tendencias generales de los datos experimentales y es compatible con el valor medido.
- Hemos explorado la dependencia con la energía de ε_2 y ε_3, codificada en el crecimiento del radio del hotspot, desde ISR hasta el LHC encontrando pequeñas desviaciones en sus valores.
- El grueso de nuestros resultados para $\varepsilon_{2(3)}$ sigue siendo cualitativamente el mismo al variar la cantidad de hot spots de tres a $N_{hs} = (2, 4)$. Se observa un gradual aumento de $\varepsilon_{2(3)}$ al incrementar N_{hs}.
- Exponemos, por primera vez en la literatura, un mecanismo concreto que permite una anti-correlación de ε_2 y ε_3 en los bines de centralidad más altos tal y como dictan los datos. Al modelar el protón como compuesto por 3 hot spots, la suposición más común en la literatura, encontramos que la inclusión de correlaciones espaciales es indispensable para reproducir el signo negativo de NSC(2,3).
- Confirmamos el gran poder de discriminación de este observable en cualquier modelo de estado inicial realista mediante un scan del espacio de parámetros. Esto da lugar a estrictas restricciones en la relación R_{hs}/r_c dentro de nuestro modelo.

- Los valores de NSC(2,3) son sensibles a la variación del número de hot spots que constituyen el protón. Específicamente, encontramos que, dentro del escenario correlacionado, agregar un hot spot adicional a nuestra descripción refuerza el signo negativo de NSC(2,3) mientras que el considerar $N_{hs} = 2$ empuja NSC(2,3) hacia valores positivos. Para $N_{hs} = 4$ obtenemos un valor negativo de NSC(2,3) dentro del escenario no correlacionado aunque su valor absoluto sigue siendo menor que en el caso correlacionado.

- Interpretamos este último resultado mostrando indicios de un comportamiento universal de NSC(2,3) cuando se evalúa como una función de N_w/N_{coll}. El peso de las configuraciones con un gran valor de esta cantidad en la simulación Monte Carlo tiene una influencia dramática en el signo de NSC(2,3). En el escenario correlacionado aumenta la probabilidad de tener configuraciones en las que muchos hot spots interaccionan un reducido número de veces. Estas topologías de interacción son responsables en nuestro modelo del signo negativo del cumulante simétrico NSC(2,3).

Los resultados presentados en esta tesis deben ser considerados como los primeros pasos hacia un objetivo más amplio: la comprensión dinámica del origen de las correlaciones espaciales dentro del protón y sus implicaciones fenomenológicas. Con el fin de profundizar en este tema hay dos grandes líneas de investigación que exploraremos en el futuro cercano.

Por un lado, para hacer una comparación unívoca de nuestro modelo de estado inicial con los datos experimentales sobre coeficientes armónicos de flujo y cumulantes simétricos alimentaremos una simulación hidrodinámica, viscosa y relativista con los perfiles de densidad de entropía iniciales obtenidos en nuestro modelo Monte Carlo Glauber. Este proyecto nos permitirá verificar si el efecto de las correlaciones espaciales desaparece tras la evolución hidrodinámica o, por el contrario, da una mejor descripción de los datos experimentales. Si se supone una relación lineal entre los momentos de excentricidad espacial y los coeficientes armónicos de flujo, se espera que el efecto neto de las correlaciones espaciales sea un incremento del flujo elíptico y triangular en las interacciones protón-protón de alta multiplicidad. Además, nos gustaría calibrar la sensibilidad de nuestros resultados al tamaño del sistema ampliando nuestro marco para incluir colisiones p+Pb. En sistemas tan pequeños, es imprescindible demostrar que las suposiciones subyacentes a la aplicabilidad de la hidrodinámica siguen siendo válidas.

Por otro lado, en aras de reforzar las bases teóricas de nuestro modelo para el protón, estamos trabajando en su realización explícita dentro del formalismo Color Glass Condensate. La introducción de correlaciones espaciales entre un número reducido de fuentes de carga de color, es decir, los hot spots, implicaría ir más allá de la aproximación tomada en el exitoso modelo McLerran-Venugopalan de una distribución de probabilidad gaussiana para las cargas de color, justificada por el teorema del límite central cuando el número de ellas es lo suficientemente grande como en el caso nuclear. Por lo tanto, el objetivo final de este estudio además de su contenido teórico es medir el impacto de un número finito de fuentes de color correlacionadas en las correlaciones de momento del estado inicial.

Completar de manera satisfactoria ambos proyectos abrirá la puerta a una mayor comprensión de los intrigantes indicios experimentales de la producción de gotas de plasma de quarks y gluones en pequeños sistemas de colisión desde una perspectiva tanto de estado inicial como de final.

A

Kinematics

A particle position and momentum are characterized in terms of four-vectors. They are defined as

$$x^{\mu} = (t, x, y, z)$$
$$p^{\mu} = (E, p_x, p_y, p_z). \tag{A.1}$$

To operate with four-vectors, a metric tensor has to be specified. In this thesis, two systems of coordinates are used: Minkowski and light-cone. In Minkowski space the metric is defined as $g^{\mu\nu} = \mathrm{diag}(1, -1, -1, -1)$. The definition of light-cone coordinates is given in Appendix B. The z-direction is consider the longitudinal one and (x, y) often collapse into a transverse component i.e. $\mathbf{x}_{\perp} \equiv \sqrt{x^2 + y^2}$.

Although in general we work with Lorentz invariant objects the physics interpretation of any process varies with the reference frame we choose to work in. In our case, they will be the centre-of-mass (C.o.M) and Breit frames. Consider a collision between particles A and B. In the C.o.M frame, the relation $\vec{p}_A + \vec{p}_B = 0$ is fulfilled. For the Breit frame, the initial momentum of A is reversed after the scattering i.e. $\vec{p}_A^{\,\mathrm{out}} = -\vec{p}_A^{\,\mathrm{in}}$.

Consider the hadronic collision between particles A and B in which a relativistic particle of mass m is created with momentum p and energy $E = \sqrt{m^2 + p^2}$. We take the z-direction as the collision axis. The angle between the transverse momentum $\mathbf{p}_{\perp}$ and the longitudinal component p_z is the polar emission angle θ. Unfortunately, p_z does not transform linearly under Lorentz transformations so a new variable called *rapidity* is introduced

$$y = \frac{1}{2} \log \left(\frac{E + p_z}{E - p_z} \right). \tag{A.2}$$

Central rapidity is equivalent to $y \sim 0$ while large values of rapidity are referred to as *forward*. Note that the definition of rapidity involves the mass of the particle. An alternative variable, the *pseudo-rapidity*, that can be extracted from the particle track geometry in the detector is often used:

$$\begin{aligned}\eta &= \frac{1}{2}\log\left(\frac{\sqrt{\mathbf{p}_\perp{}^2+p_z^2}+p_z}{\sqrt{\mathbf{p}_\perp{}^2+p_z^2}-p_z}\right)\\ &= \frac{1}{2}\log\left(\frac{1+\cos\theta}{1-\cos\theta}\right). \end{aligned} \tag{A.3}$$

In the ultra-relativistic limit $y \approx \eta$. The centre of mass energy of the collision is defined as

$$\sqrt{s} = \sqrt{2E_A E_B} \tag{A.4}$$

Further, the four-momentum transfer squared can be expressed as:

$$t = (p_A^\mu - p^\mu)^2 = -2p^2(1-\cos\theta). \tag{A.5}$$

Finally, by measuring $(y, \mathbf{p}_\perp)$ of the produced particle and knowing $\sqrt{s}$ one has access to the longitudinal momentum fraction carried by the participant partons of their parent hadrons. Assuming that the produced particle originates from a single partonic interaction, the longitudinal momentum fractions are given by

$$x_{1,2} = \frac{p_\perp}{\sqrt{s}} e^{\pm y}. \tag{A.6}$$

B

Light-cone coordinates

A coordinate system frequently used in relativistic collisions is defined by choosing an axis along which particles have a large momentum and calling it x^3. Then, for any 4-vector x^μ we define:

$$x^\mu = (x^+, x^-, \mathbf{x}_\perp) \tag{B.1}$$

with

$$x^\pm = \frac{1}{\sqrt{2}}(x^0 \pm x^3) \tag{B.2}$$

and $\mathbf{x}_\perp = (x^1, x^2)$ being (x^0, x^1, x^2, x^3) the coordinates in the standard basis of Minkowski space. The graphic representation of the two different coordinate systems is shown in Fig. B.1.

In order to obtain the metric, $g^{\mu\nu}$ of this coordinate system,

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu \quad (\text{Minkowski} \quad ds^2 = dt^2 - dx^2 - dy^2 - dz^2) \tag{B.3}$$

$$2dx^+ dx^- = (dx^0 + dx^3)(dx^0 - dx^3) = (dx^0)^2 - (dx^3)^2 \tag{B.4}$$

$$\Rightarrow ds^2_{LC} = 2dx^x dx^- - (dx^1)^2 - (dx^2)^2 \tag{B.5}$$

Assuming that we write $x^\mu = (x^+, x^-, \mathbf{x}_\perp)$, the metric looks like

$$g_{\mu\nu} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

.

A product of two four-vectors a^μ and b^μ in light cone notation is

$$a \cdot b = g_{\mu\nu} a^\nu b^\mu = a_\mu b^\mu = a^+ b^- + a^- b^+ - \mathbf{a}_\perp \cdot \mathbf{b}_\perp \tag{B.6}$$

$$a \cdot a = 2a^+ a^- - \mathbf{a}_\perp \cdot \mathbf{a}_\perp . \tag{B.7}$$

Finally, the on-shell condition is

$$k_\mu k^\mu = m^2 \Rightarrow 2k^+ k^- - k_\perp^2 = m^2 \Rightarrow k^- = \frac{k_\perp^2 + m^2}{2k^+}. \tag{B.8}$$

It is also useful to define the proper time and the space time rapidity as follows

$$\begin{aligned} \tau &= \sqrt{(x^0)^2 - (x^3)^2} = \sqrt{2x^+ x^-} \\ \eta &= \frac{1}{2} \ln \frac{x^0 + x^3}{x^0 - x^3} = \frac{1}{2} \ln \frac{x^+}{x^-} \end{aligned} \tag{B.9}$$

The metric in the $(\tau, \eta, \mathbf{x}_\perp)$ reads

$$ds^2 = d\tau^2 - \tau^2 d\eta - d\mathbf{x}_\perp \tag{B.10}$$

Therefore, any 4-vector $A^\mu(x^+, x^-, \mathbf{x}_\perp)$ can be transformed to $(\tau, \eta, \mathbf{x}_\perp)$ coordinates by applying

$$\begin{aligned} A_\tau &= A^\tau = (x^+ A^- + x^- A^+)/\tau \\ A_\eta &= -\tau^2 A^\eta = x^+ A^- - x^- A^+ . \end{aligned} \tag{B.11}$$

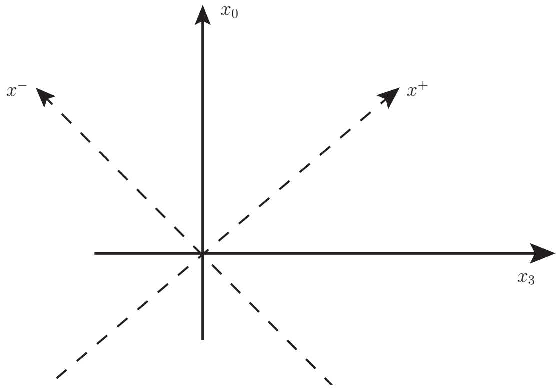

Figure B.1: Representation of the rotation from usual Minkowski space (x_0, x_3) to light-cone coordinates (x^+, x^-).

C

Eikonal scattering and Wilson lines

In Chapter 1 we have related the gauge link needed to obtain a gauge invariant definition of Wigner distributions to the rescattering of partons off their parent hadron in terms of Wilson lines. To justify this idea we present an explicit calculation of the S-matrix of the process depicted in Fig. C.1. That is, a fast parton exchange one gluon with a color field source that can be its parent hadron, a nucleus or a nucleon. Generically we will refer to it as *target*.

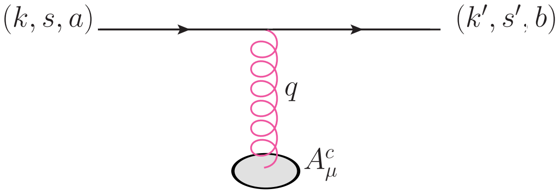

Figure C.1: One gluon exchange between a fast quark and a dense target.

Before delving into the details of the calculation we make some kinematic considerations. In the infinite momentum frame, the target moves with a very large momentum along the x^- direction i.e. $P \approx (0, P^-, \mathbf{0}_\perp)$. Therefore, the color field only has a $-$ component, $A^c_\mu \equiv A^{c,-}(x^+, \mathbf{x}_\perp)$ and it is independent of x^-. A simple $+ \leftrightarrow -$ sustitution relates the results of this Appendix and the ones in Sec. 1.3. On the other hand, the massless quark is moving along the light-cone x^+ axis with a very large momentum k^+. Applying the on-shell condition we get the magnitude of k^-

$$k^2 = 0 \Rightarrow 2k^+k^- - k_\perp^2 = 0 \Rightarrow k^- = \frac{k_\perp^2}{2k^+}. \tag{C.1}$$

Now we apply the eikonal approximation: at high energies, the recoil of the propagating quark as it scatters off the target can be neglected. In transverse space, this is equivalent to assume the transverse position of the quark to be fixed during the propagation. Consequently, $|k_\perp| \ll k^+$ and from Eq. (C.1) one sees that k^- is strongly suppressed. Then, the 4-momentum of the incoming quark is

$$k^\mu \approx (k^+, 0, \mathbf{k}_\perp) \tag{C.2}$$

.

Regarding the outgoing quark, its 4-momentum given by:

$$k' = (k^+ + q^+, q^-, \mathbf{k}_\perp + \mathbf{q}_\perp) \tag{C.3}$$

where q is the 4-momentum of the gluon. This expression can be further simplified by, first, imposing the on-shell condition

$$k'^2 = 0 \Rightarrow q^2 + 2k^+ q^- = 0 \Rightarrow q^- = -\frac{q^2}{2k^+} \to 0 \tag{C.4}$$

An alternative, but equivalent, view of the eikonal approximation is that the energy of the collision is much greater than the energy exchanged between parton and background field so that the latter can be neglected. Therefore, $p^+ \gg q^+$ and the final expression for the outgoing quark is

$$k' \approx (k^+, 0, \mathbf{k}_\perp + \mathbf{q}_\perp) \tag{C.5}$$

Now we are ready to evaluate the scattering matrix corresponding to Fig. C.1 that reads

$$S_1(k, k', a, b, s, s') = \int \mathrm{d}^4 x e^{\mathrm{i}(k-k')x} \overline{u}^{s'}(k') \mathrm{i} g A^c_\mu (t^c)_{ab} \gamma^\mu u^s(k) \tag{C.6}$$

where (a, b) are color indices in the fundamental representation and (s, s') relate to the spin. Using light-cone coordinates

$$\begin{aligned} S_1 = \mathrm{i}g \int \mathrm{d}x^+ \mathrm{d}x^- \mathrm{d}^2 x_\perp e^{\mathrm{i}(k^+ - k'^+)x^-} e^{\mathrm{i}(k^- - k'^-)x^+} e^{-\mathrm{i}(\mathbf{k}_\perp - \mathbf{k}'_\perp)\cdot \mathbf{x}_\perp} \\ \times \overline{u}^{s'}(k') \mathrm{i} g A^c_\mu (t^c)_{ab} \gamma^\mu u^s(k). \end{aligned} \tag{C.7}$$

Next, it can be shown that $\overline{u}^{s'} \gamma^\mu u^s = 2\delta^{ss'} k^\mu$ within the eikonal approximation [213] and defining $A_\mu \equiv A^c_\mu (t^c)_{ab}$ Eq. (C.7) turns into

$$S_1 = \mathrm{i}g \delta^{ss'} \int \mathrm{d}x^+ \mathrm{d}x^- \mathrm{d}^2 x_\perp e^{\mathrm{i}(k^+ - k'^+)x^-} e^{\mathrm{i}(k^- - k'^-)x^+} e^{-\mathrm{i}(\mathbf{k}_\perp - \mathbf{k}'_\perp)\cdot \mathbf{x}_\perp} 2k^\mu A_\mu. \tag{C.8}$$

where the dependence of A_μ on the light-cone coordinates has been omitted. As mentioned above, A_μ has only a $-$ component so that its 4-product with k^μ is simplified

$$S_1 = 2\mathrm{i}g \delta^{ss'} \int \mathrm{d}x^+ \mathrm{d}x^- \mathrm{d}^2 x_\perp e^{\mathrm{i}(k^+ - k'^+)x^-} e^{\mathrm{i}(k^- - k'^-)x^+} e^{-\mathrm{i}(\mathbf{k}_\perp - \mathbf{k}'_\perp)\cdot \mathbf{x}_\perp} k^+ A_-. \tag{C.9}$$

Making use of the shape of the four momenta of the incoming and outgoing quark as derived in Eqs. (C.2) and (C.5) we get:

$$\begin{aligned}S_1 &= 2\mathrm{i}g\delta^{ss'}k^+\int \mathrm{d}x^+\mathrm{d}x^-\mathrm{d}^2x_\perp e^{\mathrm{i}(k^+-k'^+)x^-}e^{-\mathrm{i}(\mathbf{k}_\perp-\mathbf{k}'_\perp)\cdot\mathbf{x}_\perp}A_-(x^+,\mathbf{x}_\perp)\\ &= 2\mathrm{i}g2\pi\delta^{ss'}k^+\delta(k^+-k'^+)\int \mathrm{d}x^+\mathrm{d}^2x_\perp e^{-\mathrm{i}(\mathbf{k}_\perp-\mathbf{k}'_\perp)\cdot\mathbf{x}_\perp}A_-(x^+,\mathbf{x}_\perp).\end{aligned} \tag{C.10}$$

To make the connection with the Wilson line clearer we restore the color matrix and write:

$$S_1 = 2\pi\delta(k^+-k'^+)\delta^{ss'}2k^+\int \mathrm{d}^2x_\perp e^{-\mathrm{i}(\mathbf{k}_\perp-\mathbf{k}'_\perp)\cdot\mathbf{x}_\perp}\left[\mathrm{i}g\int \mathrm{d}x^+A^{c-}(x^+,\mathbf{x}_\perp)t^c\right]_{ab} \tag{C.11}$$

The factors outside the integral in Eq. (C.11) reflect the conservation of the spin and the $+$-component of the momentum in the scattering. The term in brackets has a neat interpretation: the quark undergoes a color rotation from a to b due to the scattering process. Up to now we have considered only one gluon exchange i.e. Eq. (C.11) is the first term of an infinte sum that takes into account all possible scatterings. When this sum is performed, the total scattering matrix reads:

$$\begin{aligned}S &= \sum_{n=0}^{\infty} S_n\\ &= \delta(k^+-k'^+)\delta^{ss'}2k^+\int \mathrm{d}^2x_\perp e^{-\mathrm{i}(\mathbf{k}_\perp-\mathbf{k}'_\perp)\cdot\mathbf{x}_\perp}\mathcal{P}\exp\left[\mathrm{i}g\int \mathrm{d}x^+A^{c-}(x^+,\mathbf{x}_\perp)t^c\right]_{ab}.\end{aligned} \tag{C.12}$$

Finally, we can write S as:

$$S(k,k',a,b,s,s') = \delta(k^+-k'^+)\delta^{ss'}2k^+\int \mathrm{d}^2x_\perp e^{-\mathrm{i}(\mathbf{k}_\perp-\mathbf{k}'_\perp)\cdot\mathbf{x}_\perp}\mathcal{U}(\mathbf{x}_\perp)_{ab} \tag{C.13}$$

where

$$\mathcal{U}(\mathbf{x}_\perp)_{ab} = \mathcal{P}\exp\left[\mathrm{i}g\int \mathrm{d}x^+A^{c-}(x^+,\mathbf{x}_\perp)t^c\right]_{ab} \tag{C.14}$$

is the definition of a Wilson line.

List of Figures

List of Tables

Bibliography

[1] R. K. Ellis, W. J. Stirling, B. R. Webber, QCD and collider physics, Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol. 8 (1996) 1–435.

[2] M. E. Peskin, D. V. Schroeder, An Introduction to quantum field theory, Addison-Wesley, Reading, USA, 1995.
URL `http://www.slac.stanford.edu/~mpeskin/QFT.html`

[3] L. D. Faddeev, V. N. Popov, Feynman Diagrams for the Yang-Mills Field, Phys. Lett. 25B (1967) 29–30. `doi:10.1016/0370-2693(67)90067-6`.

[4] D. J. Gross, F. Wilczek, Ultraviolet Behavior of Nonabelian Gauge Theories, Phys. Rev. Lett. 30 (1973) 1343–1346. `doi:10.1103/PhysRevLett.30.1343`.

[5] V. Khachatryan, et al., Constraints on parton distribution functions and extraction of the strong coupling constant from the inclusive jet cross section in pp collisions at $\sqrt{s} = 7$ TeV, Eur. Phys. J. C75 (6) (2015) 288. `arXiv:1410.6765`, `doi:10.1140/epjc/s10052-015-3499-1`.

[6] E. P. Wigner, On the quantum correction for thermodynamic equilibrium, Phys. Rev. 40 (1932) 749–760. `doi:10.1103/PhysRev.40.749`.

[7] A. V. Belitsky, X.-d. Ji, F. Yuan, Quark imaging in the proton via quantum phase space distributions, Phys. Rev. D69 (2004) 074014. `arXiv:hep-ph/0307383`, `doi:10.1103/PhysRevD.69.074014`.

[8] S. J. Brodsky, H.-C. Pauli, S. S. Pinsky, Quantum chromodynamics and other field theories on the light cone, Phys. Rept. 301 (1998) 299–486. `arXiv:hep-ph/9705477`, `doi:10.1016/S0370-1573(97)00089-6`.

[9] C. Lorce, B. Pasquini, Quark Wigner Distributions and Orbital Angular Momentum, Phys. Rev. D84 (2011) 014015. `arXiv:1106.0139`, `doi:10.1103/PhysRevD.84.014015`.

[10] A. V. Belitsky, A. V. Radyushkin, Unraveling hadron structure with generalized parton distributions, Phys. Rept. 418 (2005) 1–387. `arXiv:hep-ph/0504030`, `doi:10.1016/j.physrep.2005.06.002`.

[11] J. C. Collins, D. E. Soper, G. F. Sterman, Factorization of Hard Processes in QCD, Adv. Ser. Direct. High Energy Phys. 5 (1989) 1–91. `arXiv:hep-ph/0409313`, `doi:10.1142/9789814503266_0001`.

[12] J. Collins, Foundations of perturbative QCD, Cambridge monographs on particle physics, nuclear physics, and cosmology, Cambridge Univ. Press, New York, NY, 2011.
URL `https://cds.cern.ch/record/1350496`

[13] R. Devenish, A. Cooper-Sarkar, Deep Inelastic Scattering, Oxford University Press, 2004.
URL `https://books.google.de/books?id=L68zn24GgkcC`

[14] I. O. Cherednikov, T. Mertens, F. F. Van der Veken, Wilson lines in quantum field theory, Vol. 24, De Gruyter, 2014.
URL `http://www.degruyter.com/view/product/204486`

[15] F. Dominguez, Unintegrated Gluon Distributions at Small-x, Ph.D. thesis, Columbia U. (2011).
URL `http://inspirehep.net/record/1280362/files/Thesis-2011-Dominguez.pdf`

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

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

[18] G. Altarelli, G. Parisi, Asymptotic Freedom in Parton Language, Nucl. Phys. B126 (1977) 298–318. `doi:10.1016/0550-3213(77)90384-4`.

[19] J. L. Albacete, J. I. Illana, A. Soto-Ontoso, Neutrino-nucleon cross section at ultrahigh energy and its astrophysical implications, Phys. Rev. D92 (1) (2015) 014027. `arXiv:1505.06583`, `doi:10.1103/PhysRevD.92.014027`.

[20] R. D. Ball, et al., Parton distributions from high-precision collider data, Eur. Phys. J. C77 (10) (2017) 663. `arXiv:1706.00428`, `doi:10.1140/epjc/s10052-017-5199-5`.

[21] A. Accardi, et al., Electron Ion Collider: The Next QCD Frontier, Eur. Phys. J. A52 (9) (2016) 268. `arXiv:1212.1701`, `doi:10.1140/epja/i2016-16268-9`.

[22] D. Sivers, Single-spin production asymmetries from the hard scattering of pointlike constituents, Phys. Rev. D 41 (1990) 83–90. `doi:10.1103/PhysRevD.41.83`.
URL `https://link.aps.org/doi/10.1103/PhysRevD.41.83`

[23] D. Boer, P. J. Mulders, Time reversal odd distribution functions in leptoproduction, Phys. Rev. D57 (1998) 5780–5786. `arXiv:hep-ph/9711485`, `doi:10.1103/PhysRevD.57.5780`.

[24] J. C. Collins, D. E. Soper, Back-To-Back Jets in QCD, Nucl. Phys. B193 (1981) 381, [Erratum: Nucl. Phys.B213,545(1983)]. `doi:10.1016/0550-3213(81)90339-4`.

[25] X.-d. Ji, J.-p. Ma, F. Yuan, QCD factorization for semi-inclusive deep-inelastic scattering at low transverse momentum, Phys. Rev. D71 (2005) 034005. `arXiv:hep-ph/0404183`, `doi:10.1103/PhysRevD.71.034005`.

[26] M. G. Echevarria, A. Idilbi, I. Scimemi, Factorization Theorem For Drell-Yan At Low q_T And Transverse Momentum Distributions On-The-Light-Cone, JHEP 07 (2012) 002. `arXiv:1111.4996`, `doi:10.1007/JHEP07(2012)002`.

[27] J. Collins, J.-W. Qiu, k_T factorization is violated in production of high-transverse-momentum particles in hadron-hadron collisions, Phys. Rev. D75 (2007) 114014. `arXiv:0705.2141`, `doi:10.1103/PhysRevD.75.114014`.

[28] A. Airapetian, et al., Observation of the Naive-T-odd Sivers Effect in Deep-Inelastic Scattering, Phys. Rev. Lett. 103 (2009) 152002. `arXiv:0906.3918`, `doi:10.1103/PhysRevLett.103.152002`.

[29] M. Alekseev, et al., Collins and Sivers asymmetries for pions and kaons in muon-deuteron DIS, Phys. Lett. B673 (2009) 127–135. `arXiv:0802.2160`, `doi:10.1016/j.physletb.2009.01.060`.

[30] X. Qian, et al., Single Spin Asymmetries in Charged Pion Production from Semi-Inclusive Deep Inelastic Scattering on a Transversely Polarized ^{3}He Target, Phys. Rev. Lett. 107 (2011) 072003. `arXiv:1106.0363`, `doi:10.1103/PhysRevLett.107.072003`.

[31] M. Aghasyan, et al., First measurement of transverse-spin-dependent azimuthal asymmetries in the Drell-Yan process, Phys. Rev. Lett. 119 (11) (2017) 112002. `arXiv:1704.00488`, `doi:10.1103/PhysRevLett.119.112002`.

[32] L. Adamczyk, et al., Measurement of the transverse single-spin asymmetry in $p^{\uparrow} + p \to W^{\pm}/Z^0$ at RHIC, Phys. Rev. Lett. 116 (13) (2016) 132301. `arXiv:1511.06003`, `doi:10.1103/PhysRevLett.116.132301`.

[33] C. Adolph, et al., II – Experimental investigation of transverse spin asymmetries in µ -p SIDIS processes: Sivers asymmetries, Phys. Lett. B717 (2012) 383–389. `arXiv:1205.5122`, `doi:10.1016/j.physletb.2012.09.056`.

[34] J. C. Collins, A. Freund, Proof of factorization for deeply virtual Compton scattering in QCD, Phys. Rev. D59 (1999) 074009. `arXiv:hep-ph/9801262`, `doi:10.1103/PhysRevD.59.074009`.

[35] M. Diehl, Generalized parton distributions, Phys. Rept. 388 (2003) 41–277. `arXiv:hep-ph/0307382`, `doi:10.1016/j.physrep.2003.08.002,10.3204/DESY-THESIS-2003-018`.

[36] T. DeGrand, C. E. Detar, Lattice methods for quantum chromodynamics, 2006.

[37] K. G. Wilson, Confinement of Quarks, Phys. Rev. D10 (1974) 2445–2459, [,45(1974)]. `doi:10.1103/PhysRevD.10.2445`.

[38] P. Hagler, Hadron structure from lattice quantum chromodynamics, Phys. Rept. 490 (2010) 49–175. `arXiv:0912.5483`, `doi:10.1016/j.physrep.2009.12.008`.

[39] B. U. Musch, P. Hagler, M. Engelhardt, J. W. Negele, A. Schafer, Sivers and Boer-Mulders observables from lattice QCD, Phys. Rev. D85 (2012) 094510. `arXiv:1111.4249`, `doi:10.1103/PhysRevD.85.094510`.

[40] M. Gockeler, R. Horsley, D. Pleiter, P. E. L. Rakow, G. Schierholz, A Lattice determination of moments of unpolarised nucleon structure functions using improved Wilson fermions, Phys. Rev. D71 (2005) 114511. `arXiv:hep-ph/0410187`, `doi:10.1103/PhysRevD.71.114511`.

[41] D. Dolgov, et al., Moments of nucleon light cone quark distributions calculated in full lattice QCD, Phys. Rev. D66 (2002) 034506. `arXiv:hep-lat/0201021`, `doi:10.1103/PhysRevD.66.034506`.

[42] E. Iancu, R. Venugopalan, The Color glass condensate and high-energy scattering in QCD, in: In *Hwa, R.C. (ed.) et al.: Quark gluon plasma* 249-3363, 2003. `arXiv:hep-ph/0303204`, `doi:10.1142/9789812795533_0005`.

[43] F. Gelis, E. Iancu, J. Jalilian-Marian, R. Venugopalan, The Color Glass Condensate, Ann. Rev. Nucl. Part. Sci. 60 (2010) 463–489. `arXiv:1002.0333`, `doi:10.1146/annurev.nucl.010909.083629`.

[44] I. Balitsky, A. Tarasov, Rapidity evolution of gluon TMD from low to moderate x, JHEP 10 (2015) 017. `arXiv:1505.02151`, `doi:10.1007/JHEP10(2015)017`.

[45] I. Balitsky, Operator expansion for high-energy scattering, Nucl. Phys. B463 (1996) 99–160. `arXiv:hep-ph/9509348`, `doi:10.1016/0550-3213(95)00638-9`.

[46] Y. V. Kovchegov, Small x F(2) structure function of a nucleus including multiple pomeron exchanges, Phys. Rev. D60 (1999) 034008. `arXiv:hep-ph/9901281`, `doi:10.1103/PhysRevD.60.034008`.

[47] M. Aaboud, et al., Measurement of the total cross section from elastic scattering in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector, Phys. Lett. B761 (2016) 158–178. `arXiv:1607.06605`, `doi:10.1016/j.physletb.2016.08.020`.

[48] J. L. Albacete, C. Marquet, Gluon saturation and initial conditions for relativistic heavy ion collisions, Prog. Part. Nucl. Phys. 76 (2014) 1–42. `arXiv:1401.4866`, `doi:10.1016/j.ppnp.2014.01.004`.

[49] L. D. McLerran, R. Venugopalan, Computing quark and gluon distribution functions for very large nuclei, Phys. Rev. D49 (1994) 2233–2241. `arXiv:hep-ph/9309289`, `doi:10.1103/PhysRevD.49.2233`.

[50] J. Jalilian-Marian, A. Kovner, L. D. McLerran, H. Weigert, The Intrinsic glue distribution at very small x, Phys. Rev. D55 (1997) 5414–5428. `arXiv:hep-ph/9606337`, `doi:10.1103/PhysRevD.55.5414`.

[51] M. Froissart, Asymptotic behavior and subtractions in the Mandelstam representation, Phys. Rev. 123 (1961) 1053–1057. `doi:10.1103/PhysRev.123.1053`.

[52] A. Kovner, U. A. Wiedemann, Nonlinear QCD evolution: Saturation without unitarization, Phys. Rev. D66 (2002) 051502. `arXiv:hep-ph/0112140`, `doi:10.1103/PhysRevD.66.051502`.

[53] K. J. Golec-Biernat, A. M. Stasto, On solutions of the Balitsky-Kovchegov equation with impact parameter, Nucl. Phys. B668 (2003) 345–363. `arXiv:hep-ph/0306279`, `doi:10.1016/j.nuclphysb.2003.07.011`.

[54] J. L. Albacete, N. Armesto, J. G. Milhano, P. Quiroga-Arias, C. A. Salgado, AAMQS: A non-linear QCD analysis of new HERA data at small-x including heavy quarks, Eur. Phys. J. C71 (2011) 1705. `arXiv:1012.4408`, `doi:10.1140/epjc/s10052-011-1705-3`.

[55] K. Aamodt, et al., Centrality dependence of the charged-particle multiplicity density at mid-rapidity in Pb-Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV, Phys. Rev. Lett. 106 (2011) 032301. `arXiv:1012.1657`, `doi:10.1103/PhysRevLett.106.032301`.

[56] C. Marquet, E. Petreska, C. Roiesnel, Transverse-momentum-dependent gluon distributions from JIMWLK evolution, JHEP 10 (2016) 065. `arXiv:1608.02577`, `doi:10.1007/JHEP10(2016)065`.

[57] G. Antchev, et al., Proton-proton elastic scattering at the LHC energy of s** (1/2) = 7-TeV, EPL 95 (4) (2011) 41001. `arXiv:1110.1385`, `doi:10.1209/0295-5075/95/41001`.

[58] E. Ruiz Arriola, W. Broniowski, Proton–Proton On Shell Optical Potential at High Energies and the Hollowness Effect, Few Body Syst. 57 (7) (2016) 485–490. `arXiv:1602.00288`, `doi:10.1007/s00601-016-1095-z`.

[59] J. L. Albacete, A. Soto-Ontoso, Hot spots and the hollowness of proton–proton interactions at high energies, Phys. Lett. B770 (2017) 149–153. `arXiv:1605.09176`, `doi:10.1016/j.physletb.2017.04.055`.

[60] M. M. Block, Hadronic forward scattering: Predictions for the Large Hadron Collider and cosmic rays, Phys. Rept. 436 (2006) 71–215. `arXiv:hep-ph/0606215`, `doi:10.1016/j.physrep.2006.06.003`.

[61] J. R. Cudell, V. V. Ezhela, P. Gauron, K. Kang, Yu. V. Kuyanov, S. B. Lugovsky, E. Martynov, B. Nicolescu, E. A. Razuvaev, N. P. Tkachenko, Benchmarks for the forward observables at RHIC, the Tevatron Run II and the LHC, Phys. Rev. Lett. 89 (2002) 201801. `arXiv:hep-ph/0206172`, `doi:10.1103/PhysRevLett.89.201801`.

[62] G. Antchev, et al., Measurement of elastic pp scattering at $\sqrt{\mathrm{s}} = 8$ TeV in the Coulomb–nuclear interference region: determination of the ρ -parameter and the total cross-section, Eur. Phys. J. C76 (12) (2016) 661. `arXiv:1610.00603`, `doi:10.1140/epjc/s10052-016-4399-8`.

[63] W. Broniowski, E. Ruiz Arriola, Hollowness in pp scattering at the LHC, Acta Phys. Polon. Supp. 10 (2017) 1203. `arXiv:1708.00402`, `doi:10.5506/APhysPolBSupp.10.1203`.

[64] J. Prochazka, V. Kundrat, Eikonal model analysis of elastic hadron collisions at high energies`arXiv:1606.09479`.

[65] J. Kaspar, Elastic scattering at the LHC, Ph.D. thesis, Prague, Inst. Phys. (2011). URL `http://inspirehep.net/record/1186254/files/CERN-THESIS-2011-214.pdf`

[66] G. Arnison, et al., Elastic and total cross-section measurement at the CERN proton-antiproton collider, Phys. Lett. 128B (1983) 336, [,15(1983)]. `doi:10.1016/0370-2693(83)90271-X`.

[67] N. A. Amos, et al., $\bar{p}p$ elastic scattering at $\sqrt{s}$ = 1.8-TeV from |t| = $0.034 - GeV/c^2$ to $0.65 - GeV/c^2$, Phys. Lett. B247 (1990) 127–130. `doi:10.1016/0370-2693(90)91060-O`.

[68] U. Amaldi, K. R. Schubert, Impact Parameter Interpretation of Proton Proton Scattering from a Critical Review of All ISR Data, Nucl. Phys. B166 (1980) 301–320. `doi:10.1016/0550-3213(80)90229-1`.

[69] T. Csörgő, Evidence for Non-Exponential Differential Cross-Section of pp Elastic Scattering at Low $|t|$ and $\sqrt{s}{=}8$ TeV by TOTEM, EPJ Web Conf. 120 (2016) 02004. `arXiv:1602.00219`, `doi:10.1051/epjconf/201612002004`.

[70] A. Donnachie, P. V. Landshoff, Elastic Scattering at Large t, Z. Phys. C2 (1979) 55, [Erratum: Z. Phys.C2,372(1979)]. `doi:10.1007/BF01546237`.

[71] C. Ewerz, The Odderon in quantum chromodynamics`arXiv:hep-ph/0306137`.

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

[73] E. Ruiz Arriola, W. Broniowski, Proton-proton hollowness at the LHC from inverse scattering, Phys. Rev. D95 (7) (2017) 074030. `arXiv:1609.05597`, `doi:10.1103/PhysRevD.95.074030`.

[74] A. Alkin, E. Martynov, O. Kovalenko, S. M. Troshin, Impact-parameter analysis of TOTEM data at the LHC: Black disk limit exceeded, Phys. Rev. D89 (9) (2014) 091501. `arXiv:1403.8036`, `doi:10.1103/PhysRevD.89.091501`.

[75] I. M. Dremin, Will protons become gray at 13 TeV and 100 TeV?, Bull. Lebedev Phys. Inst. 44 (4) (2017) 94–98. `arXiv:1511.03212`, `doi:10.3103/S1068335617040029`.

[76] S. M. Troshin, N. E. Tyurin, The new scattering mode emerging at the LHC?, Mod. Phys. Lett. A31 (13) (2016) 1650079. `arXiv:1602.08972`, `doi:10.1142/S0217732316500796`.

[77] F. James, M. Roos, MINUIT: a system for function minimization and analysis of the parameter errors and corrections, Comput. Phys. Commun. 10 (CERN-DD-75-20) (1975) 343–367. 38 p.
URL `https://cds.cern.ch/record/310399`

[78] I. M. Dremin, A toy model of elastic scattering of high energy protons`arXiv:1702.06304`.

[79] I. M. Dremin, Unexpected properties of interaction of high-energy protons, Phys. Usp. 60 (4) (2017) 333–344. `arXiv:1610.07937`, `doi:10.3367/UFNe.2016.11.037977`.

[80] V. A. Petrov, A. P. Samokhin, Is There a Hollow Inside the Proton?, in: 31th International Workshop on High Energy Physics: Critical points in the modern particle physics (IHEP2017) Protvino, Russia, July 5-7, 2017, 2018. `arXiv:1801.03809`.
URL `http://inspirehep.net/record/1647568/files/arXiv:1801.03809.pdf`

[81] W. Broniowski, E. Ruiz Arriola, Hollowness in pp scattering, Acta Phys. Polon. B48 (2017) 927. `arXiv:1704.03271`, `doi:10.5506/APhysPolB.48.927`.

[82] A. Di Giacomo, H. Panagopoulos, Field strength correlations in the QCD vacuum, Phys. Lett. B285 (1992) 133–136. `doi:10.1016/0370-2693(92)91311-V`.

[83] E. V. Shuryak, I. Zahed, Understanding the nonperturbative deep inelastic scattering: Instanton induced inelastic dipole dipole cross-section, Phys. Rev. D69 (2004) 014011. `arXiv:hep-ph/0307103`, `doi:10.1103/PhysRevD.69.014011`.

[84] T. Schäfer, E. V. Shuryak, Instantons in QCD, Rev. Mod. Phys. 70 (1998) 323–426. `arXiv:hep-ph/9610451`, `doi:10.1103/RevModPhys.70.323`.

[85] V. M. Braun, P. Gornicki, L. Mankiewicz, A. Schafer, Gluon form-factor of the proton from QCD sum rules, Phys. Lett. B302 (1993) 291–298. `doi:10.1016/0370-2693(93)90399-3`.

[86] B. Z. Kopeliovich, I. K. Potashnikova, B. Povh, E. Predazzi, Soft QCD dynamics of elastic scattering in impact parameter representation, Phys. Rev. D63 (2001) 054001. `arXiv:hep-ph/0009008`, `doi:10.1103/PhysRevD.63.054001`.

[87] A. Kovner, U. A. Wiedemann, Perturbative saturation and the soft pomeron, Phys. Rev. D66 (2002) 034031. `arXiv:hep-ph/0204277`, `doi:10.1103/PhysRevD.66.034031`.

[88] S. Bondarenko, E. Levin, J. Nyiri, Recent experimental data and the size of the quark in the constituent quark model, Eur. Phys. J. C25 (2002) 277–286. `arXiv:hep-ph/0204156`, `doi:10.1140/s10052-002-0996-9`.

[89] S. Domdey, H.-J. Pirner, U. A. Wiedemann, Testing the Scale Dependence of the Scale Factor sigma(eff) in Double Dijet Production at the LHC, Eur. Phys. J. C65 (2010) 153–162. `arXiv:0906.4335`, `doi:10.1140/epjc/s10052-009-1198-5`.

[90] R. J. Glauber, G. Matthiae, High-energy scattering of protons by nuclei, Nucl. Phys. B21 (1970) 135–157. `doi:10.1016/0550-3213(70)90511-0`.

[91] A. Bialas, K. Fialkowski, W. Slominski, M. Zielinski, Elastic p p Cross-Section and Multiple Scattering of Quarks, Acta Phys. Polon. B8 (1977) 855–874.

[92] M. Alvioli, H. J. Drescher, M. Strikman, A Monte Carlo generator of nucleon configurations in complex nuclei including Nucleon-Nucleon correlations, Phys. Lett. B680 (2009) 225–230. `arXiv:0905.2670`, `doi:10.1016/j.physletb.2009.08.067`.

[93] J.-P. Blaizot, W. Broniowski, J.-Y. Ollitrault, Correlations in the Monte Carlo Glauber model, Phys. Rev. C90 (3) (2014) 034906. `arXiv:1405.3274`, `doi:10.1103/PhysRevC.90.034906`.

[94] C. Bierlich, G. Gustafson, L. Lönnblad, A. Tarasov, Effects of Overlapping Strings in pp Collisions, JHEP 03 (2015) 148. `arXiv:1412.6259`, `doi:10.1007/JHEP03(2015)148`.

[95] Wolfram Research, Inc., Mathematica 8.0.
URL `https://www.wolfram.com`

[96] J. C. Collins, M. J. Perry, Superdense Matter: Neutrons Or Asymptotically Free Quarks?, Phys. Rev. Lett. 34 (1975) 1353. `doi:10.1103/PhysRevLett.34.1353`.

[97] D. H. Rischke, The Quark gluon plasma in equilibrium, Prog. Part. Nucl. Phys. 52 (2004) 197–296. `arXiv:nucl-th/0305030`, `doi:10.1016/j.ppnp.2003.09.002`.

[98] A. Bazavov, et al., The QCD Equation of State to $\mathcal{O}(\mu_B^6)$ from Lattice QCD, Phys. Rev. D95 (5) (2017) 054504. `arXiv:1701.04325`, `doi:10.1103/PhysRevD.95.054504`.

[99] O. Philipsen, The QCD equation of state from the lattice, Prog. Part. Nucl. Phys. 70 (2013) 55–107. `arXiv:1207.5999`, `doi:10.1016/j.ppnp.2012.09.003`.

[100] H.-T. Ding, F. Karsch, S. Mukherjee, Thermodynamics of strong-interaction matter from Lattice QCD, Int. J. Mod. Phys. E24 (10) (2015) 1530007. `arXiv:1504.05274`, `doi:10.1142/S0218301315300076`.

[101] A. Bazavov, et al., Equation of state in (2+1)-flavor QCD, Phys. Rev. D90 (2014) 094503. `arXiv:1407.6387`, `doi:10.1103/PhysRevD.90.094503`.

[102] S. Borsanyi, Z. Fodor, C. Hoelbling, S. D. Katz, S. Krieg, K. K. Szabo, Full result for the QCD equation of state with 2+1 flavors, Phys. Lett. B730 (2014) 99–104. `arXiv:1309.5258`, `doi:10.1016/j.physletb.2014.01.007`.

[103] G. Baym, T. Hatsuda, T. Kojo, P. D. Powell, Y. Song, T. Takatsuka, From hadrons to quarks in neutron stars: a review`arXiv:1707.04966`.

[104] M. Hanauske, J. Steinheimer, L. Bovard, A. Mukherjee, S. Schramm, K. Takami, J. Papenfort, N. Wechselberger, L. Rezzolla, H. Stöcker, Concluding Remarks: Connecting Relativistic Heavy Ion Collisions and Neutron Star Mergers by the Equation of State of Dense Hadron- and Quark Matter as signalled by Gravitational Waves, J. Phys. Conf. Ser. 878 (1) (2017) 012031. `doi:10.1088/1742-6596/878/1/012031`.

[105] K. Takami, L. Rezzolla, L. Baiotti, Constraining the Equation of State of Neutron Stars from Binary Mergers, Phys. Rev. Lett. 113 (9) (2014) 091104. `arXiv:1403.5672`, `doi:10.1103/PhysRevLett.113.091104`.

[106] B. Abbott, et al., GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral, Phys. Rev. Lett. 119 (16) (2017) 161101. `arXiv:1710.05832`, `doi:10.1103/PhysRevLett.119.161101`.

[107] Y. Aoki, G. Endrodi, Z. Fodor, S. D. Katz, K. K. Szabo, The Order of the quantum chromodynamics transition predicted by the standard model of particle physics, Nature 443 (2006) 675–678. `arXiv:hep-lat/0611014`, `doi:10.1038/nature05120`.

[108] A. Bzdak, V. Koch, N. Strodthoff, Cumulants and correlation functions versus the QCD phase diagram, Phys. Rev. C95 (5) (2017) 054906. `arXiv:1607.07375`, `doi:10.1103/PhysRevC.95.054906`.

[109] A. Bashir, L. Chang, I. C. Cloet, B. El-Bennich, Y.-X. Liu, C. D. Roberts, P. C. Tandy, Collective perspective on advances in Dyson-Schwinger Equation QCD, Commun. Theor. Phys. 58 (2012) 79–134. `arXiv:1201.3366`, `doi:10.1088/0253-6102/58/1/16`.

[110] T. Ablyazimov, et al., Challenges in QCD matter physics –The scientific programme of the Compressed Baryonic Matter experiment at FAIR, Eur. Phys. J. A53 (3) (2017) 60. `arXiv:1607.01487`, `doi:10.1140/epja/i2017-12248-y`.

[111] E. C. Aschenauer, S. Fazio, M. A. C. Lamont, H. Paukkunen, P. Zurita, Nuclear Structure Functions at a Future Electron-Ion Collider, Phys. Rev. D96 (11) (2017) 114005. `arXiv:1708.05654`, `doi:10.1103/PhysRevD.96.114005`.

[112] J. L. Abelleira Fernandez, et al., A Large Hadron Electron Collider at CERN, in: CERN Council Open Symposium on European Strategy for Particle Physics (ESPP2012) Cracow, Poland, September 10-12, 2012, 2012. `arXiv:1211.4831`. URL `https://inspirehep.net/record/1203309/files/arXiv:1211.4831.pdf`

[113] T. Lappi, L. McLerran, Some features of the glasma, Nucl. Phys. A772 (2006) 200–212. `arXiv:hep-ph/0602189`, `doi:10.1016/j.nuclphysa.2006.04.001`.

[114] A. Kovner, L. D. McLerran, H. Weigert, Gluon production from nonAbelian Weizsacker-Williams fields in nucleus-nucleus collisions, Phys. Rev. D52 (1995) 6231–6237. `arXiv:hep-ph/9502289`, `doi:10.1103/PhysRevD.52.6231`.

[115] T. Lappi, Production of gluons in the classical field model for heavy ion collisions, Phys. Rev. C 67 (2003) 054903. `doi:10.1103/PhysRevC.67.054903`.
URL `https://link.aps.org/doi/10.1103/PhysRevC.67.054903`

[116] T. Lappi, Energy density of the glasma, Phys. Lett. B643 (2006) 11–16. `arXiv:hep-ph/0606207`, `doi:10.1016/j.physletb.2006.10.017`.

[117] T. Epelbaum, F. Gelis, Pressure isotropization in high energy heavy ion collisions, Phys. Rev. Lett. 111 (2013) 232301. `arXiv:1307.2214`, `doi:10.1103/PhysRevLett.111.232301`.

[118] K. Fukushima, F. Gelis, The evolving Glasma, Nucl. Phys. A874 (2012) 108–129. `arXiv:1106.1396`, `doi:10.1016/j.nuclphysa.2011.11.003`.

[119] F. Gelis, Initial State and Thermalization in the Color Glass Condensate Framework, in: X.-N. Wang (Ed.), Quark-Gluon Plasma 5, 2016, pp. 67–129. `doi:10.1142/9789814663717_0002`.

[120] R. Baier, A. H. Mueller, D. Schiff, D. T. Son, 'Bottom up' thermalization in heavy ion collisions, Phys. Lett. B502 (2001) 51–58. `arXiv:hep-ph/0009237`, `doi:10.1016/S0370-2693(01)00191-5`.

[121] A. Kurkela, Y. Zhu, Isotropization and hydrodynamization in weakly coupled heavy-ion collisions, Phys. Rev. Lett. 115 (18) (2015) 182301. `arXiv:1506.06647`, `doi:10.1103/PhysRevLett.115.182301`.

[122] L. Keegan, A. Kurkela, A. Mazeliauskas, D. Teaney, Initial conditions for hydrodynamics from weakly coupled pre-equilibrium evolution, JHEP 08 (2016) 171. `arXiv:1605.04287`, `doi:10.1007/JHEP08(2016)171`.

[123] M. P. Heller, R. A. Janik, P. Witaszczyk, The characteristics of thermalization of boost-invariant plasma from holography, Phys. Rev. Lett. 108 (2012) 201602. `arXiv:1103.3452`, `doi:10.1103/PhysRevLett.108.201602`.

[124] P. M. Chesler, L. G. Yaffe, Holography and colliding gravitational shock waves in asymptotically AdS_5 spacetime, Phys. Rev. Lett. 106 (2011) 021601. `arXiv:1011.3562`, `doi:10.1103/PhysRevLett.106.021601`.

[125] A. Mazeliauskas, Talk presented at Initial Stages 2017, Krakow (Poland), September 18-22, 2017.

[126] D. Kharzeev, M. Nardi, Hadron production in nuclear collisions at RHIC and high density QCD, Phys. Lett. B507 (2001) 121–128. `arXiv:nucl-th/0012025`, `doi:10.1016/S0370-2693(01)00457-9`.

[127] J. L. ALbacete, A. Dumitru, A model for gluon production in heavy-ion collisions at the LHC with rcBK unintegrated gluon densities`arXiv:1011.5161`.

[128] B. Schenke, P. Tribedy, R. Venugopalan, Fluctuating Glasma initial conditions and flow in heavy ion collisions, Phys. Rev. Lett. 108 (2012) 252301. `arXiv:1202.6646`, `doi:10.1103/PhysRevLett.108.252301`.

[129] B. Alver, M. Baker, C. Loizides, P. Steinberg, The PHOBOS Glauber Monte Carlo`arXiv:0805.4411`.

[130] H. Kowalski, D. Teaney, An Impact parameter dipole saturation model, Phys. Rev. D68 (2003) 114005. `arXiv:hep-ph/0304189`, `doi:10.1103/PhysRevD.68.114005`.

[131] D. H. Rischke, Fluid dynamics for relativistic nuclear collisions[Lect. Notes Phys.516,21(1999)]. `arXiv:nucl-th/9809044`, `doi:10.1007/BFb0107310`.

[132] W. A. Hiscock, L. Lindblom, Stability and causality in dissipative relativistic fluids, Annals Phys. 151 (1983) 466–496. `doi:10.1016/0003-4916(83)90288-9`.

[133] W. Israel, J. M. Stewart, Transient relativistic thermodynamics and kinetic theory, Annals Phys. 118 (1979) 341–372. `doi:10.1016/0003-4916(79)90130-1`.

[134] G. S. Denicol, H. Niemi, E. Molnar, D. H. Rischke, Derivation of transient relativistic fluid dynamics from the Boltzmann equation, Phys. Rev. D85 (2012) 114047, [Erratum: Phys. Rev.D91,no.3,039902(2015)]. `arXiv:1202.4551`, `doi:10.1103/PhysRevD.85.114047,10.1103/PhysRevD.91.039902`.

[135] M. Alqahtani, M. Nopoush, M. Strickland, Relativistic anisotropic hydrodynamics`arXiv:1712.03282`.

[136] G. S. Denicol, E. Molnár, H. Niemi, D. H. Rischke, Derivation of fluid dynamics from kinetic theory with the 14-moment approximation, Eur. Phys. J. A48 (2012) 170. `arXiv:1206.1554`, `doi:10.1140/epja/i2012-12170-x`.

[137] P. Romatschke, M. Strickland, Collective modes of an anisotropic quark gluon plasma, Phys. Rev. D68 (2003) 036004. `arXiv:hep-ph/0304092`, `doi:10.1103/PhysRevD.68.036004`.

[138] M. Alqahtani, M. Nopoush, R. Ryblewski, M. Strickland, Anisotropic hydrodynamic modeling of 2.76 TeV Pb-Pb collisions, Phys. Rev. C96 (4) (2017) 044910. `arXiv:1705.10191`, `doi:10.1103/PhysRevC.96.044910`.

[139] H. Niemi, G. S. Denicol, How large is the Knudsen number reached in fluid dynamical simulations of ultrarelativistic heavy ion collisions?`arXiv:1404.7327`.

[140] F. Cooper, G. Frye, E. Schonberg, Landau's Hydrodynamic Model of Particle Production and electron Positron Annihilation Into Hadrons, Phys. Rev. D11 (1975) 192. `doi:10.1103/PhysRevD.11.192`.

[141] H. Petersen, Anisotropic flow in transport + hydrodynamics hybrid approaches, J. Phys. G41 (12) (2014) 124005. `arXiv:1404.1763`, `doi:10.1088/0954-3899/41/12/124005`.

[142] S. A. Bass, et al., Microscopic models for ultrarelativistic heavy ion collisions, Prog. Part. Nucl. Phys. 41 (1998) 255–369, [Prog. Part. Nucl. Phys.41,225(1998)]. `arXiv:nucl-th/9803035`, `doi:10.1016/S0146-6410(98)00058-1`.

[143] J. Weil, et al., Particle production and equilibrium properties within a new hadron transport approach for heavy-ion collisions, Phys. Rev. C94 (5) (2016) 054905. `arXiv:1606.06642`, `doi:10.1103/PhysRevC.94.054905`.

[144] U. A. Wiedemann, Jet Quenching in Heavy Ion Collisions (2010) 521–562[Landolt-Bornstein23,521(2010)]. `arXiv:0908.2306`, `doi:10.1007/978-3-642-01539-7_17`.

[145] J. D. Bjorken, Energy Loss of Energetic Partons in Quark - Gluon Plasma: Possible Extinction of High p(t) Jets in Hadron - Hadron Collisions.

[146] B. Abelev, et al., Transverse momentum distribution and nuclear modification factor of charged particles in p-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV, Phys. Rev. Lett. 110 (8) (2013) 082302. `arXiv:1210.4520`, `doi:10.1103/PhysRevLett.110.082302`.

[147] V. Khachatryan, et al., Charged-particle nuclear modification factors in PbPb and pPb collisions at $\sqrt{s_{\mathrm{N\,N}}} = 5.02$ TeV, JHEP 04 (2017) 039. `arXiv:1611.01664`, `doi:10.1007/JHEP04(2017)039`.

[148] M. J. Tannenbaum, Critical examination of RHIC paradigms: Mostly high p_T, PoS CERP2010 (2010) 019. `arXiv:1008.1536`.

[149] J.-Y. Ollitrault, Anisotropy as a signature of transverse collective flow, Phys. Rev. D46 (1992) 229–245. `doi:10.1103/PhysRevD.46.229`.

[150] A. M. Poskanzer, S. A. Voloshin, Methods for analyzing anisotropic flow in relativistic nuclear collisions, Phys. Rev. C58 (1998) 1671–1678. `arXiv:nucl-ex/9805001`, `doi:10.1103/PhysRevC.58.1671`.

[151] A. Adams, L. D. Carr, T. Schaefer, P. Steinberg, J. E. Thomas, Focus on strongly correlated quantum fluids: From ultracold quantum gases to QCD plasmas, New J. Phys. 15 (2013) 045022. `doi:10.1088/1367-2630/15/045022`.

[152] S. S. Adler, et al., Elliptic flow of identified hadrons in Au+Au collisions at s(NN)**(1/2) = 200-GeV, Phys. Rev. Lett. 91 (2003) 182301. `arXiv:nucl-ex/0305013`, `doi:10.1103/PhysRevLett.91.182301`.

[153] P. Huovinen, P. F. Kolb, U. W. Heinz, P. V. Ruuskanen, S. A. Voloshin, Radial and elliptic flow at RHIC: Further predictions, Phys. Lett. B503 (2001) 58–64. `arXiv:hep-ph/0101136`, `doi:10.1016/S0370-2693(01)00219-2`.

[154] C. Gale, S. Jeon, B. Schenke, P. Tribedy, R. Venugopalan, Event-by-event anisotropic flow in heavy-ion collisions from combined Yang-Mills and viscous fluid dynamics, Phys. Rev. Lett. 110 (1) (2013) 012302. `arXiv:1209.6330`, `doi:10.1103/PhysRevLett.110.012302`.

[155] V. Khachatryan, et al., Observation of Long-Range Near-Side Angular Correlations in Proton-Proton Collisions at the LHC, JHEP 09 (2010) 091. `arXiv:1009.4122`, `doi:10.1007/JHEP09(2010)091`.

[156] N. Borghini, P. M. Dinh, J.-Y. Ollitrault, Are flow measurements at SPS reliable?, Phys. Rev. C62 (2000) 034902. `arXiv:nucl-th/0004026`, `doi:10.1103/PhysRevC.62.034902`.

[157] N. Borghini, P. M. Dinh, J.-Y. Ollitrault, Flow analysis from multiparticle azimuthal correlations, Phys. Rev. C 64 (2001) 054901. `doi:10.1103/PhysRevC.64.054901`. URL `https://link.aps.org/doi/10.1103/PhysRevC.64.054901`

[158] R. S. Bhalerao, N. Borghini, J. Y. Ollitrault, Analysis of anisotropic flow with Lee-Yang zeroes, Nucl. Phys. A727 (2003) 373–426. `arXiv:nucl-th/0310016`, `doi:10.1016/j.nuclphysa.2003.08.007`.

[159] A. Bilandzic, C. H. Christensen, K. Gulbrandsen, A. Hansen, Y. Zhou, Generic framework for anisotropic flow analyses with multiparticle azimuthal correlations, Phys. Rev. C89 (6) (2014) 064904. `arXiv:1312.3572`, `doi:10.1103/PhysRevC.89.064904`.

[160] M. Aaboud, et al., Measurement of multi-particle azimuthal correlations in pp, p+Pb and low-multiplicity Pb+Pb collisions with the ATLAS detector, Eur. Phys. J. C77 (6) (2017) 428. `arXiv:1705.04176`, `doi:10.1140/epjc/s10052-017-4988-1`.

[161] Z.-W. Lin, C. M. Ko, B.-A. Li, B. Zhang, S. Pal, A Multi-phase transport model for relativistic heavy ion collisions, Phys. Rev. C72 (2005) 064901. `arXiv:nucl-th/0411110`, `doi:10.1103/PhysRevC.72.064901`.

[162] R. Snellings, Elliptic Flow: A Brief Review, New J. Phys. 13 (2011) 055008. `arXiv:1102.3010`, `doi:10.1088/1367-2630/13/5/055008`.

[163] C. Adler, et al., Elliptic flow from two and four particle correlations in Au+Au collisions at s(NN)**(1/2) = 130-GeV, Phys. Rev. C66 (2002) 034904. `arXiv:nucl-ex/0206001`, `doi:10.1103/PhysRevC.66.034904`.

[164] S. A. Voloshin, A. M. Poskanzer, R. Snellings, Collective phenomena in non-central nuclear collisions`arXiv:0809.2949`.

[165] B. B. Abelev, et al., Multiparticle azimuthal correlations in p -Pb and Pb-Pb collisions at the CERN Large Hadron Collider, Phys. Rev. C90 (5) (2014) 054901. `arXiv:1406.2474`, `doi:10.1103/PhysRevC.90.054901`.

[166] K. Gajdošová, Investigations of anisotropic collectivity using multi-particle correlations in pp, p–Pb and Pb–Pb collisions, Nucl. Phys. A967 (2017) 437–440. `doi:10.1016/j.nuclphysa.2017.04.033`.

[167] J. Jia, M. Zhou, A. Trzupek, Revealing long-range multiparticle collectivity in small collision systems via subevent cumulants, Phys. Rev. C 96 (2017) 034906. `doi:10.1103/PhysRevC.96.034906`.
URL `https://link.aps.org/doi/10.1103/PhysRevC.96.034906`

[168] V. Khachatryan, et al., Evidence for collectivity in pp collisions at the LHC, Phys. Lett. B765 (2017) 193–220. `arXiv:1606.06198`, `doi:10.1016/j.physletb.2016.12.009`.

[169] G. Giacalone, L. Yan, J. Noronha-Hostler, J.-Y. Ollitrault, Symmetric cumulants and event-plane correlations in Pb + Pb collisions, Phys. Rev. C94 (1) (2016) 014906. `arXiv:1605.08303`, `doi:10.1103/PhysRevC.94.014906`.

[170] A. M. Sirunyan, et al., Observation of correlated azimuthal anisotropy Fourier harmonics in pp and pPb collisions at the LHC`arXiv:1709.09189`.

[171] J. Adam, et al., Correlated event-by-event fluctuations of flow harmonics in Pb-Pb collisions at $\sqrt{s_{\rm NN}} = 2.76$ TeV, Phys. Rev. Lett. 117 (2016) 182301. `arXiv:1604.07663`, `doi:10.1103/PhysRevLett.117.182301`.

[172] J. Adam, et al., Enhanced production of multi-strange hadrons in high-multiplicity proton-proton collisions, Nature Phys. 13 (2017) 535–539. `arXiv:1606.07424`, `doi:10.1038/nphys4111`.

[173] P. Bozek, W. Broniowski, Correlations from hydrodynamic flow in p-Pb collisions, Phys. Lett. B718 (2013) 1557–1561. `arXiv:1211.0845`, `doi:10.1016/j.physletb.2012.12.051`.

[174] H. Niemi, K. J. Eskola, R. Paatelainen, Event-by-event fluctuations in a perturbative QCD + saturation + hydrodynamics model: Determining QCD matter shear viscosity in ultrarelativistic heavy-ion collisions, Phys. Rev. C93 (2) (2016) 024907. `arXiv:1505.02677`, `doi:10.1103/PhysRevC.93.024907`.

[175] R. D. Weller, P. Romatschke, One fluid to rule them all: viscous hydrodynamic description of event-by-event central p+p, p+Pb and Pb+Pb collisions at $\sqrt{s} = 5.02$ TeV, Phys. Lett. B774 (2017) 351–356. `arXiv:1701.07145`, `doi:10.1016/j.physletb.2017.09.077`.

[176] P. Bozek, W. Broniowski, Collective dynamics in high-energy proton-nucleus collisions, Phys. Rev. C88 (1) (2013) 014903. `arXiv:1304.3044`, `doi:10.1103/PhysRevC.88.014903`.

[177] K. Werner, M. Bleicher, B. Guiot, I. Karpenko, T. Pierog, Evidence for Flow from Hydrodynamic Simulations of p-Pb Collisions at 5.02 TeV from ν_2 Mass Splitting, Phys. Rev. Lett. 112 (23) (2014) 232301. `arXiv:1307.4379`, `doi:10.1103/PhysRevLett.112.232301`.

[178] F. G. Gardim, F. Grassi, M. Luzum, J. Noronha-Hostler, Hydrodynamic Predictions for Mixed Harmonic Correlations in 200 GeV Au+Au Collisions, Phys. Rev. C95 (3) (2017) 034901. `arXiv:1608.02982`, `doi:10.1103/PhysRevC.95.034901`.

[179] P. Romatschke, Do nuclear collisions create a locally equilibrated quark–gluon plasma?, Eur. Phys. J. C77 (1) (2017) 21. `arXiv:1609.02820`, `doi:10.1140/epjc/s10052-016-4567-x`.

[180] T. Lappi, B. Schenke, S. Schlichting, R. Venugopalan, Tracing the origin of azimuthal gluon correlations in the color glass condensate, JHEP 01 (2016) 061. `arXiv:1509.03499`, `doi:10.1007/JHEP01(2016)061`.

[181] A. Dumitru, K. Dusling, F. Gelis, J. Jalilian-Marian, T. Lappi, R. Venugopalan, The Ridge in proton-proton collisions at the LHC, Phys. Lett. B697 (2011) 21–25. `arXiv:1009.5295`, `doi:10.1016/j.physletb.2011.01.024`.

[182] A. Dumitru, L. McLerran, V. Skokov, Azimuthal asymmetries and the emergence of "collectivity" from multi-particle correlations in high-energy pA collisions, Phys. Lett. B743 (2015) 134–137. `arXiv:1410.4844`, `doi:10.1016/j.physletb.2015.02.046`.

[183] K. Dusling, R. Venugopalan, Azimuthal collimation of long range rapidity correlations by strong color fields in high multiplicity hadron-hadron collisions, Phys. Rev. Lett. 108 (2012) 262001. `arXiv:1201.2658`, `doi:10.1103/PhysRevLett.108.262001`.

[184] K. Dusling, M. Mace, R. Venugopalan, Parton model description of multiparticle azimuthal correlations in pA collisions, Phys. Rev. D97 (1) (2018) 016014. `arXiv:1706.06260`, `doi:10.1103/PhysRevD.97.016014`.

[185] K. Dusling, M. Mace, R. Venugopalan, Multiparticle collectivity from initial state correlations in high energy proton-nucleus collisions, Phys. Rev. Lett. 120 (4) (2018) 042002. `arXiv:1705.00745`, `doi:10.1103/PhysRevLett.120.042002`.

[186] B. Schenke, S. Schlichting, P. Tribedy, R. Venugopalan, Mass ordering of spectra from fragmentation of saturated gluon states in high multiplicity proton-proton collisions, Phys. Rev. Lett. 117 (16) (2016) 162301. `arXiv:1607.02496`, `doi:10.1103/PhysRevLett.117.162301`.

[187] M. Greif, C. Greiner, B. Schenke, S. Schlichting, Z. Xu, Importance of initial and final state effects for azimuthal correlations in p+Pb collisions, Phys. Rev. D96 (9) (2017) 091504. `arXiv:1708.02076`, `doi:10.1103/PhysRevD.96.091504`.

[188] H. Mäntysaari, B. Schenke, C. Shen, P. Tribedy, Imprints of fluctuating proton shapes on flow in proton-lead collisions at the LHC, Phys. Lett. B772 (2017) 681–686. `arXiv:1705.03177`, `doi:10.1016/j.physletb.2017.07.038`.

[189] H. Mäntysaari, B. Schenke, Evidence of strong proton shape fluctuations from incoherent diffraction, Phys. Rev. Lett. 117 (5) (2016) 052301. `arXiv:1603.04349`, `doi:10.1103/PhysRevLett.117.052301`.

[190] S. Schlichting, B. Schenke, The shape of the proton at high energies, Phys. Lett. B739 (2014) 313–319. `arXiv:1407.8458`, `doi:10.1016/j.physletb.2014.10.068`.

[191] C. Loizides, Glauber modeling of high-energy nuclear collisions at the subnucleon level, Phys. Rev. C94 (2) (2016) 024914. `arXiv:1603.07375`, `doi:10.1103/PhysRevC.94.024914`.

[192] P. Bożek, W. Broniowski, M. Rybczyński, Wounded quarks in A+A, p+A, and p+p collisions, Phys. Rev. C94 (1) (2016) 014902. `arXiv:1604.07697`, `doi:10.1103/PhysRevC.94.014902`.

[193] K. Welsh, J. Singer, U. W. Heinz, Initial state fluctuations in collisions between light and heavy ions, Phys. Rev. C94 (2) (2016) 024919. `arXiv:1605.09418`, `doi:10.1103/PhysRevC.94.024919`.

[194] J. Noronha-Hostler, L. Yan, F. G. Gardim, J.-Y. Ollitrault, Linear and cubic response to the initial eccentricity in heavy-ion collisions, Phys. Rev. C93 (1) (2016) 014909. `arXiv:1511.03896`, `doi:10.1103/PhysRevC.93.014909`.

[195] L. Yan, J.-Y. Ollitrault, $\nu_4, \nu_5, \nu_6, \nu_7$: nonlinear hydrodynamic response versus LHC data, Phys. Lett. B744 (2015) 82–87. `arXiv:1502.02502`, `doi:10.1016/j.physletb.2015.03.040`.

[196] J. L. Albacete, H. Petersen, A. Soto-Ontoso, Correlated wounded hot spots in proton-proton interactions, Phys. Rev. C95 (6) (2017) 064909. `arXiv:1612.06274`, `doi:10.1103/PhysRevC.95.064909`.

[197] J. L. Albacete, H. Petersen, A. Soto-Ontoso, Gluonic hot spots and spatial correlations inside the proton, Nucl. Phys. A967 (2017) 924–927. `doi:10.1016/j.nuclphysa.2017.05.021`.

[198] J. L. Albacete, H. Petersen, A. Soto-Ontoso, Symmetric cumulants as a probe of the proton substructure at LHC energies, Phys. Lett. B778 (2018) 128–136. `arXiv:1707.05592`, `doi:10.1016/j.physletb.2018.01.011`.

[199] R. Brun, F. Rademakers, ROOT: An object oriented data analysis framework, Nucl. Instrum. Meth. A389 (1997) 81–86. `doi:10.1016/S0168-9002(97)00048-X`.

[200] W. Broniowski, M. Rybczynski, P. Bozek, GLISSANDO: Glauber initial-state simulation and more.., Comput. Phys. Commun. 180 (2009) 69–83. `arXiv:0710.5731`, `doi:10.1016/j.cpc.2008.07.016`.

[201] J. E. Bernhard, J. S. Moreland, S. A. Bass, J. Liu, U. Heinz, Applying Bayesian parameter estimation to relativistic heavy-ion collisions: simultaneous characterization of the initial state and quark-gluon plasma medium, Phys. Rev. C94 (2) (2016) 024907. `arXiv:1605.03954`, `doi:10.1103/PhysRevC.94.024907`.

[202] J. T. Mitchell, D. V. Perepelitsa, M. J. Tannenbaum, P. W. Stankus, Tests of constituent-quark generation methods which maintain both the nucleon center of mass and the desired radial distribution in Monte Carlo Glauber models, Phys. Rev. C93 (5) (2016) 054910. `arXiv:1603.08836`, `doi:10.1103/PhysRevC.93.054910`.

[203] A. Bialas, W. Czyz, W. Furmanski, Particle Production in Hadron-Nucleus Collisions and the Quark Model, Acta Phys. Polon. B8 (1977) 585.

[204] A. Bialas, M. Bleszynski, W. Czyz, Multiplicity Distributions in Nucleus-Nucleus Collisions at High-Energies, Nucl. Phys. B111 (1976) 461–476. `doi:10.1016/0550-3213(76)90329-1`.

[205] J. Adam, et al., Charged-particle multiplicities in proton–proton collisions at $\sqrt{s} = 0.9$ to 8 TeV, Eur. Phys. J. C77 (1) (2017) 33. `arXiv:1509.07541`, `doi:10.1140/epjc/s10052-016-4571-1`.

[206] M. Aaboud, et al., Charged-particle distributions at low transverse momentum in $\sqrt{s} = 13$ TeV pp interactions measured with the ATLAS detector at the LHC, Eur. Phys. J. C76 (9) (2016) 502. `arXiv:1606.01133`, `doi:10.1140/epjc/s10052-016-4335-y`.

[207] K. Aamodt, et al., Charged-particle multiplicity measurement in proton-proton collisions at $\sqrt{s} = 7$ TeV with ALICE at LHC, Eur. Phys. J. C68 (2010) 345–354. `arXiv:1004.3514`, `doi:10.1140/epjc/s10052-010-1350-2`.

[208] A. Breakstone, et al., Charged Multiplicity Distribution in p p Interactions at ISR Energies, Phys. Rev. D30 (1984) 528. `doi:10.1103/PhysRevD.30.528`.

[209] G. S. Denicol, C. Gale, S. Jeon, J. F. Paquet, B. Schenke, Effect of initial-state nucleon-nucleon correlations on collective flow in ultra-central heavy-ion collisions`arXiv:1406.7792`.

[210] G. Giacalone, J. Noronha-Hostler, J.-Y. Ollitrault, Relative flow fluctuations as a probe of initial state fluctuations, Phys. Rev. C95 (5) (2017) 054910. `arXiv:1702.01730`, `doi:10.1103/PhysRevC.95.054910`.

[211] H. Petersen, Identified Particle Spectra and Anisotropic Flow in an Event-by-Event Hybrid Approach in Pb+Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 2.76$ TeV, Phys. Rev. C84 (2011) 034912. `arXiv:1105.1766`, `doi:10.1103/PhysRevC.84.034912`.

[212] W. Broniowski, P. Bozek, M. Rybczynski, Wounded quarks at the LHC, Acta Phys. Polon. Supp. 10 (2017) 513. `arXiv:1611.00250`, `doi:10.5506/APhysPolBSupp.10.513`.

[213] G. P. Lepage, S. J. Brodsky, Exclusive Processes in Perturbative Quantum Chromodynamics, Phys. Rev. D22 (1980) 2157. `doi:10.1103/PhysRevD.22.2157`.