Photon tagged jet substructure measurements in pp and PbPb collisions with the CMS detector

Bharadwaj Harikrishnan

INSTITUT POLYTECHNIQUE DE PARIS

NNT : 2024IPPAX076

Thèse de doctorat

Photon-tagged jet substructure measurements in pp and PbPb collisions with the CMS detector

Thèse de doctorat de l'Institut Polytechnique de Paris
préparée à lÉcole polytechnique

École doctorale n°626 Ecole doctorale de lÍnstitut Polytechnique de Paris (EDIPP)
Spécialité de doctorat : Physique des particules

Thèse présentée et soutenue à Palaiseau, le 16 septembre 2024, par

BHARADWAJ HARIKRISHNAN

Composition du Jury :

Yves Sirois Directeur de recherche (École polytechnique)	President
Carlos Munoz Camacho Directeur de recherche (IJCLab)	Rapporteur
Gustavo Conesa Balbastre Chargé de recherche (LPSC)	Rapporteur
Paul Caucal Maître de conférences (Subatech)	Examinateur
Emilien Chapon Chercheur (CEA)	Examinateur
Matthew Nguyen Chargé de recherche (École polytechnique)	Directeur de thèse
Leticia Cunqueiro Mendez Professor (Sapienza Universita di Roma)	Co-directrice de thèse

To my family,

for their unwavering support,

and in memory of my late grandfather,

whose wisdom and love continues to inspire me.

Acknowledgements

First, I would like to express my deepest gratitude to my supervisors, Prof. Matthew Nguyen and Prof. Leticia Cunqueiro Mendez for their unwavering support, invaluable guidance and patience throughout my PhD journey. Thank you for giving me this opportunity. Your mentorship has been the cornerstone of my academic growth, pushing me to challenge my limits and develop as a researcher. I would also like to thank the members of my thesis committee for their feedback and for their careful reading of this document.

I am also deeply grateful to my colleagues at Laboratoire Leprince-Ringuet (LLR) for their help in all matters, academic or otherwise. Special thanks to the Compact Muon Solenoid (CMS) heavy ion group including Florian, Lida, Batoul and Guillaume for all the random office discussions, their insightful advice, and for sharing this journey with me. I am especially grateful to Cristian for sharing his expertise and his willingness to guide me through all the challenging moments in the course of my work. I would like to extend my gratitude to the Heavy Ion Physics Analysis Group (HIN PAG) in the CMS experiment whose resources and guidance has been of the utmost importance. To the High pT Physics Interest Group (PInG) in particular, I would like to thank everyone for their constructive feedback and comments which has enriched my research experience and their insights were pivotal in navigating the complexities of my research topic. I would also like to thank Thomas Hebekkar, Loukas Gouskas, Yen-Jie Lee and Stephen Mrenna for the detailed discussions and reviewing the checks for the robust analysis as

part of the Analysis Review Committee.

Finally, I would like to extend my heartfelt thanks to my family for their unconditional love allowing me to pursue my passion. The support from my parents, grandparents, uncles, aunts and cousins throughout this very important time in my life meant a lot to me. To my late grandfather, whose memory has been a guiding light, I dedicate this milestone to you. This achievement would not have been possible without the unwavering encouragement, patience, and sacrifices of my loved ones.

Contents

Chapter 1

Introduction

1.1 The Standard Model

The Standard Model (SM) of particle physics is the most successful framework of describing elementary particles and their interactions. It describes three of the four fundamental forces: electromagnetic, weak and strong force but it does not describe gravity. The description of gravity provided by general relativity has contradictions with quantum mechanics which is the basis for the SM. The effect of gravity in microscopic scale is weak enough that it is essentially unmeasureable.

Broadly, the interactions between the particles can be classified through the mediators of each fundamental force, the gauge bosons, namely,

- The photon is the mediator for the electromagnetic force and is responsible for the interactions between charged particles. The photon is massless and does not carry electric charge but is described as an excitation of the electromagnetic field in the interaction of charged particles. Electromagnetism is the only long-range force described by the SM, essentially having infinite range. The quantum counterpart to classical electromagnetism is known as Quantum Electrodynamics (QED).
- The charged $W^{\pm}$ and neutral Z^0 bosons mediate the weak interaction be-

tween all fermions. They are responsible for radioactivity and subatomic interactions. The $W^{\pm}$ bosons carry an electric charge to couple with the electromagnetic interaction as well. The $W^{\pm}$ and Z^0 bosons along with photons are collectively mediating a unified electroweak interaction.

- Gluons are the mediators for the strong force through interactions of color charges which binds the quarks to each other. There are 3 color charges, named red, blue, and green with quarks consisting of a single color charge and gluons a combination of a color and anti-color charge for a total of 8 different gluons. Since gluons contain color charges they can interact through the strong interaction with other gluons. These interactions are described by Quantum Chromodynamics (QCD)

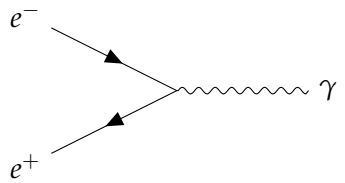


Figure 1.1: The interaction vertex in QED

The electromagnetic force is the dominant force in the interaction of atoms and nuclei studied from ancient times. Since the 18th century, a mathematical formulation of electromagnetism was developed culminating with the discovery of Maxwell's equations for a classical description of the electromagnetic field. The modern field of quantum electrodynamics (QED) is a modification of Maxwell's equations to be consistent with our understanding of the quantized nature of matter. The SM interactions are built upon the concept of gauge invariance, which states that, the equations of motion are the same upon a special set of transformations known as "gauge transformations". These transformations change the phase of the fields locally. The requirement that the equations of motion or the Lagrangian density ($\mathcal{L}$) be invariant under such transformations, requires the presence of "gauge fields" and their interactions with the SM particles.

Mathematically, QED is a U(1) symmetry group in the Abelian gauge theory with the dynamics of charged particles described by the QED Lagrangian density,

$$\mathcal{L}_{\text{QED}} = \bar{\psi}\left(i\gamma^{\mu}\partial_{\mu} - m\right)\psi - e\bar{\psi}\gamma^{\mu}A_{\mu}\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{1.1}$$

where $\psi(\bar{\psi})$ are the fields representing the charged spin $1/2$ particles involved in the interaction, γ^{μ} are Dirac matrices, ∂_{μ} is the four-gradient and A_{μ} is the gauge field of the photon which represents the the U(1) symmetry. The electromagnetic field tensor is expressed as,

$$F_{\mu\nu} = \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu} \tag{1.2}$$

$F_{\mu\nu}$ represents a kinetic term and is gauge invariant on its own. Gauge invariance, as mentioned earlier, is a property which allows for symmetric transformation of a conserved quantity. In the case of the QED Lagrangian, the charge density is a conserved quantity. As a consequence of gauge invariance in the SM, the photon field is massless, since the inclusion of a mass term of the form $m^2 A_{\mu}A^{\mu}$ for the gauge field would violate gauge invariance in U(1). The interaction of the photon with charged fermions is represented by the fundamental QED vertex shown in Fig. 1.1.

The weak interaction is a fundamental force of nuclear dynamics, responsible for radioactive decay of atoms and participates in nuclear fission and fusion. A unified electroweak theory provides a combined description of the electromagnetic and weak interactions. This is due to their combined interaction above the vacuum expectation value of 246 GeV [1]. The energy scale at which the electromagnetic and weak interactions can be unified corresponds to a temperature around 10^{15}K, present in the first picosecond after the Big Bang. The expectation value arises due to another piece of the stan-

dard model puzzle which was recently discovered in the form of the Higgs boson[2]. The presence of the Higgs boson is vital in explaining the mass of the gauge bosons (W^+, W^-, Z^0) through electroweak symmetry breaking. In the SM, at high enough temperatures all particles are massless. At this temperature, the Higgs field develops a vacuum expectation value which generates the mass for the gauge bosons such that the mass of $W^\pm \approx 80.377\,\text{GeV}$ and $Z \approx 91.1876\,\text{GeV}$.

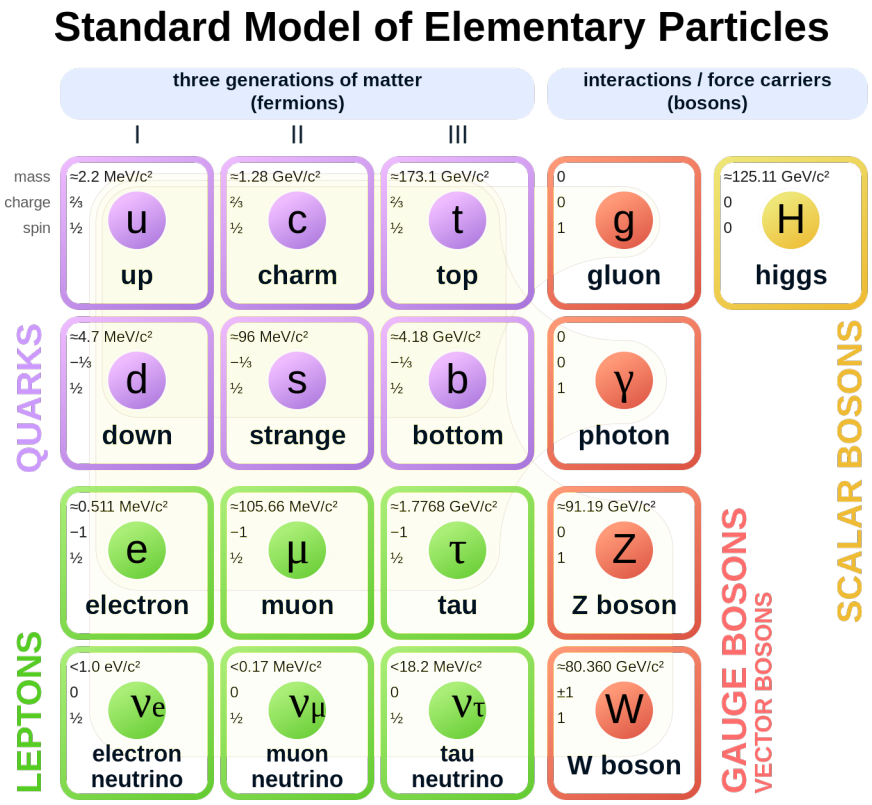


Figure 1.2: Standard Model of Particle physics [3]

The interaction of fermions with the $W^\pm$ bosons involves the change in electric charge and spin by one unit, while the Z^0 boson being neutral and being its own antiparticle cannot change the charge, and only affects the spin alongside the momentum and energy. This is called the weak interaction due to the field strength being several orders of magnitude smaller than the electromagnetic force and is consistent with the short

range due to the massive W and Z bosons. The neutrinos, which are electrically neutral and with very low rest mass, primarily interact through the weak interaction.

All fermions, quarks and leptons are divided into 3 families or generations of particles as shown in Fig. 1.2. The difference between generations are their masses and flavor quantum numbers while having the same net electric and color charge. The generations consist of a mass hierarchy causing the higher mass particles to decay to the first generation. Thus, most stable matter is made of the first generation of particles.

The focus of this thesis is on the study of the SM through hadronic collisions by analyzing the data from the Compact Muon Solenoid (CMS) experiment at the Large Hadron Collider (LHC). The collisions involving the nuclear constituents provide a window to study the strong interactions and the principles of QCD which is discussed in the following sections.

1.2 Quantum Chromodynamics (QCD)

1.2.1 The strong interaction

QCD is the fundamental field theory which describes the strong nuclear interaction. The "chromo" in QCD refers to the color charge which is analogous to electric charge in QED. Particles with the property of color charge participate through the strong interaction. On the other hand, color-neutral particles are not directly sensitive to the strong force. The strong interaction describes the fermionic components: quarks and antiquarks interacting through the mediator, gluons. Quarks are spin-1/2 particles that carry an electric charge of $-1/3$ or $+2/3$. The antiquarks consequently have a charge of $+1/3$ or $-2/3$. There are six flavors of quarks in 3 families ordered by their mass as shown in Fig. 1.2 namely, up, down, charm, strange, top, and bottom. The up(down), charm (strange) and top (bottom) quarks have the same charge of $|2/3|(|1/3|)$ and are in different generations of increasing masses. Gluons on the other hand are massless spin-1 gauge bosons. Quarks are characterized by a single color charge, while gluons are characterized by one color charge and one anti-color charge. Quarks and gluons, collectively known as partons, are known as the fundamental degrees of freedom of QCD.

There are three fundamental characteristics of QCD:

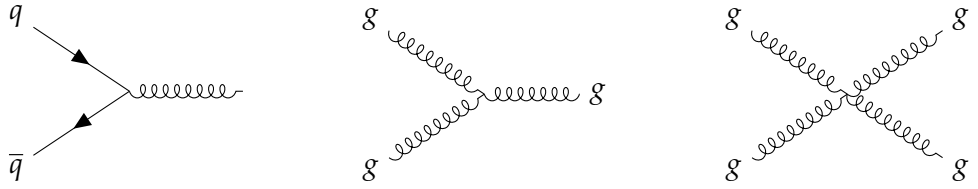
Color confinement: Individually, color charged particles are not observed directly in normal conditions. An analytic proof for this phenomenon does not exist yet and is only understood qualitatively. Since gluons themselves carry color charge, the 8 colored states of the gluon can participate in the strong interaction creating a gluon self-coupling which does not exist for photons in QED. The electric field between charged particles decreases rapidly with distance while the strong force maintains a constant field. Thus, with increasing distance, it is energetically favorable for a quark-antiquark pair to appear instead of extending the gluon field. This results in the formation of

color-neutral composite particles known as hadrons, consisting of quarks and gluons.

Asymptotic freedom: The strength of interactions between quarks and gluons decreases at short distances. Asymptotic freedom is the complementary aspect of color confinement due to the coupling between the quarks and gluons. The partonic interaction is weak at high energies allowing perturbative calculations with the varying strength of the coupling.

Chiral symmetry breaking: The primary source of the mass of matter is a result of the binding energy keeping nucleons bound together. The spontaneous symmetry breaking of a global symmetry in the light-quark sector, known as chiral symmetry, results in the generation of mass for hadrons composed of light-quarks far larger than the masses of the individual valence quarks.

1.2.2 The QCD Lagrangian



(a) Quark-Gluon Vertex (b) Three-Gluon Vertex (c) Four-Gluon Vertex

Figure 1.3: The QCD interaction vertices between quarks and gluons. Since gluons have color charge and can self-interact, in addition to the quark-gluon vertex (left) there exists three-gluon (middle) and four-gluon (right) vertices

The dynamics of the quarks and gluons are defined by the QCD Lagrangian,

$$\mathcal{L} = \bar{\psi}_i \left(i\gamma^\mu \partial_\mu - m_i \right) \psi_i - g G^\alpha_\mu \bar{\psi}_i \gamma^\mu T^\alpha_{ij} \psi_j - \frac{1}{4} G^\alpha_{\mu\nu} G^{\mu\nu}_\alpha \tag{1.3}$$

which is closely analogous to the Lagrangian of QED, the theory which governs atomic interactions. In both cases the field ψ represents a spin $1/2$ fermion which for QCD is the quark. The field G is the massless spin 1 boson field of the gluon which is coupled

to the fermion field with strength g. The gluon field tensor is expressed in terms of the gluon field as

$$G^{\alpha}_{\mu\nu} = \partial_{\mu} G^{\alpha}_{\nu} - \partial_{\nu} G^{\alpha}_{\mu} + C^{\alpha}_{\beta\gamma} G^{\beta}_{\mu} G^{\gamma}_{\nu} \tag{1.4}$$

$G^{\alpha}_{\mu\nu}$ is analogous to $F_{\mu\nu}$ in the QED Lagrangian in equation 1.1 except for the last term which represents the self interaction of the gluon. The interaction of the quarks and gluons is represented by the three QCD vertices shown in Fig. 1.3. As mentioned earlier, the gluons unlike photons in QED, carry color charge and this allows them to interact with other gluons through the strong interaction. This term is responsible for the different nature of matter at sub-nuclear scales (10^{-16} cm) compared to the atomic scale (10^{-8} cm) as shown in Fig. 1.4. The fundamental symmetries of QCD are described mathematically by the SU(3) group. Physically, this group corresponds to a theory in which there is a quantum number which can take three values, called color charge, the QCD analog to electric charge. Flavor, another quantum number of the theory, gives rise to six different known quark states (labeled by the subscript i in the Lagrangian) each with a mass and (fractional) electric charge. The three lightest quarks, up, down and strange, also obey approximate SU(3) symmetry.

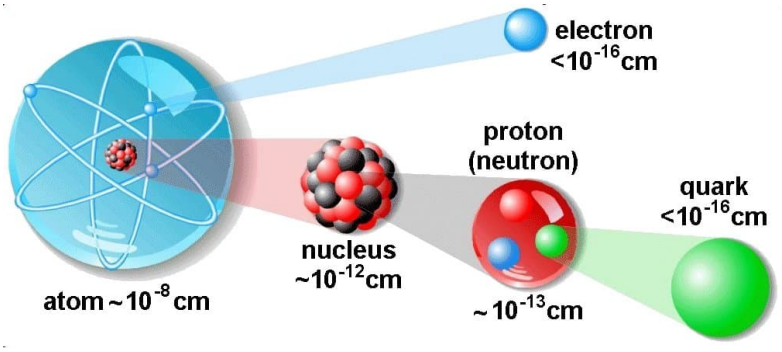


Figure 1.4: Schematic showing the scale of atomic, nuclear and subnuclear particles. Atoms have a radii of order 10^{-8} cm with nuclei of order 10^{-12} cm. The elementary particles such as electrons or quarks which form the nuclei are only of order 10^{-16} cm [4]

A key aspect of the QCD Lagrangian shown in equation 1.3 is the constant g which sets the scale of the fundamental coupling strength between quarks and gluons. The coupling strength g is not a constant and varies according to the energy scale of the interaction. The bare coupling strength g is not directly observed in nature, instead the running coupling (α_s) is measured as shown in Fig. 1.5. The running coupling goes to infinity as the scale Q^2 decreases, with the divergent scale $\Lambda_{\mathrm{QCD}} \approx 200\mathrm{MeV}$ at which perturbation theory breaks down. It also means that the theory becomes asymptotically free as Q^2 increases, or equivalently that the strong force becomes weaker at large energies, which is the opposite behavior of QED. It is therefore necessary to specify the energy scale (usually the mass of the Z boson $\approx$ 91.1876 GeV) for experimentally derived values of α_s. At leading order in QCD, α_s is determined as,

$$\alpha_s(Q^2) = \frac{g(Q^2)}{4\pi} = \frac{1}{b\log\left(Q^2/\Lambda^2_{\mathrm{QCD}}\right)} \tag{1.5}$$

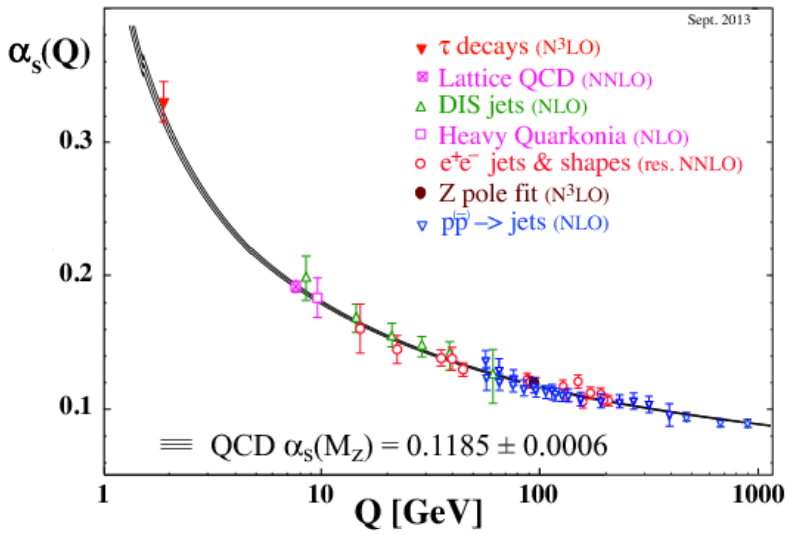


Figure 1.5: A summary of measurements of α_s as a function of energy Q^2. The measurements are compared at various orders of perturbation theory from NLO to N^{3}LO as shown to arrive at a value of $\alpha_s(M_Z) = 0.1185 \pm 0.0006$. It also illustrates the running of the strong coupling such that it goes to infinity as the scale Q^2 decreases [5]

As the coupling approaches unity, for example when the distance between two quarks

grows large, the energy carried by the fields exceeds the threshold for creation of new matter. The result is the phenomenon known as confinement in which free quarks are not observed in nature, but rather, are trapped in bound states for which the net color is zero (color singlet).

To understand the strong interaction, over the past few decades we have studied the scattering of a beam of particles with respect to fixed target or another beam of particles. The energy of the interaction determines the distances scales that can be resolved, leading to higher energy collisions being studied in recent years.

1.2.3 QCD Factorization theorem

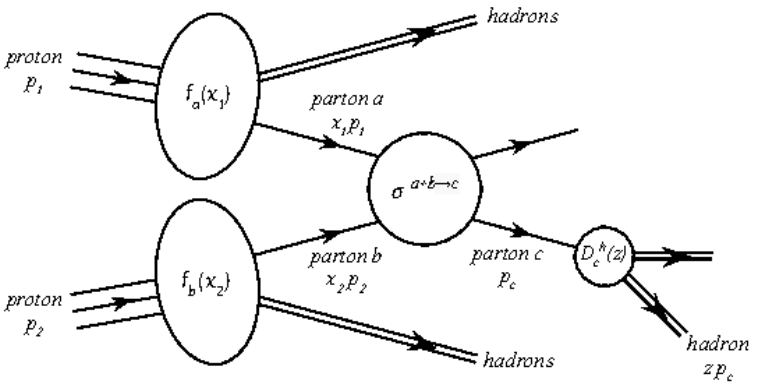


Figure 1.6: Schematic diagram of the pQCD Factorization theorem showing the parton distribution functions $f_{a,b}(x)$, partonic cross sections $\hat{\sigma} \equiv \sigma^{a+b\to c}$ and Fragmentation functions $D_c^h(z)$. It shows the incident protons p_1 and p_2 interacting to produce final state hadrons. The Bjorken x describes the fractional momentum of the parton inside the proton which participates in the interaction. [6]

In hadronic collisions, we consider the interaction to be between one parton from each of the participating nuclei. Hadrons are subatomic particles composed of (anti-)quarks and gluons held together by the strong force. When the momentum scale of the partonic interaction is much larger than Λ_{QCD}, the interaction is considered "hard" and perturbative QCD (pQCD) calculations can be used to make predictions about the interaction. However, as individual partons are not observed in nature due to color con-

finement, experimental observables need to be defined at hadron level. One of the observables is the hadron-level cross section which quantifies the production rate of physical quantities. The inclusive cross-section (σ) can be written as a convolution of the partonic cross-section ($\hat{\sigma}$), which is calculable with pQCD at short distance scales, and a parametrization of the non-calculable long range behavior called the Fragmentation Function (FF) denoted by (D_c^h). A nonperturbative component called the Parton Distribution Function (PDF) ,is necessary to complete the picture of described using the fractional momentum of the parton inside the nuclei (Bjorken x). The cross section for hadronic production in proton-proton collisions is a sum over all flavors for initial partons (a, b) to produce a parton c consisting of the products of the aforementioned contributions is given by,

$$d\sigma = \sum_{a,b,c} \int dx_a\, dx_b\, dz\, f_a(x_a)\, f_b(x_b)\, d\hat{\sigma}(p_a, p_b, p_c)\, D_c^h(z) \tag{1.6}$$

This process of separating the different components of the cross section is known as the QCD Factorization theorem [7, 8] shown schematically in Fig. 1.6. The factorization theorem holds up to corrections of $\mathcal{O}(\Lambda^2_{\mathrm{QCD}}/Q^2)$, which for hard processes ($\Lambda^2_{\mathrm{QCD}} \gg Q^2$) is negligible.

This description of the cross section comes by treating the parton distribution and fragmentation functions as independent of the species of particles involved in the collision. Through a property known as universality, it is assumed that the long range behavior of PDFs and FFs can be treated independent of the collision even though a measured cross-section does not uniquely determine these parameters.

High energy partons interacting in hadronic collisions are not observed in isolation as they lose energy through a cascade of emissions. The theoretical framework for this cascade can be understood using pQCD up to the limit when the particles created are

sufficiently soft (momentum $\approx \Lambda_{QCD}$). At this soft scale, partons are confined to color singlet states through a nonperturbative process of hadronization. This physics process of parton showering, hadronization followed by a decay is shown in Fig. 1.7.

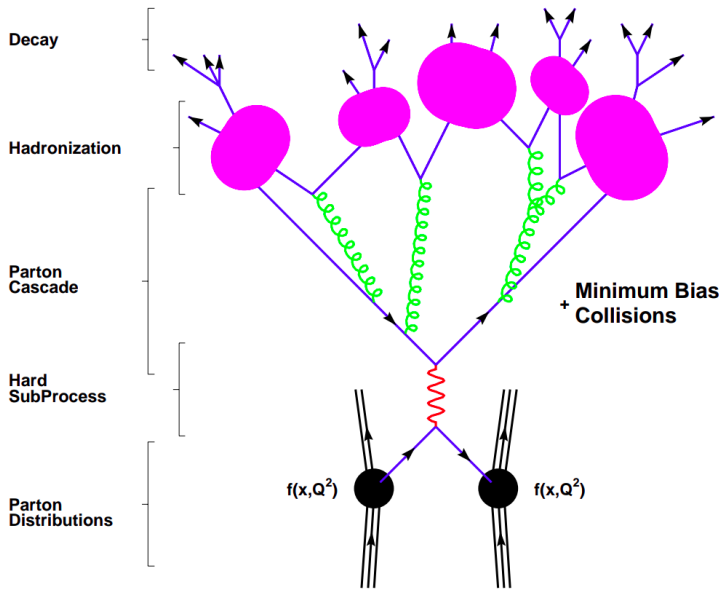


Figure 1.7: Schematic diagram of a collision event showing the parton shower and hadronization process. Incident hadrons interact considering their PDFs ($f(x, Q^2)$) participating in a hard subprocess (involving a large momentum transfer) to produce a cascade of gluons and hadronize [9]

1.2.4 Parton shower and hadronization

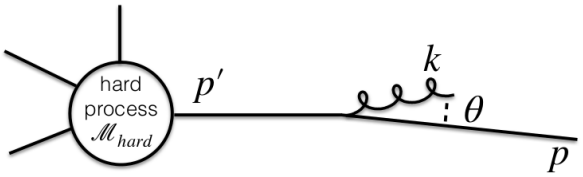


Figure 1.8: Schematic of a gluon emission from a hard process. Figure extracted from Gavin Salam's lecture notes on QCD [10]

The Fig. 1.8 above represents the single gluon emission off a parton produced in a hard elementary interaction. It represents a hard process with amplitude ($|M_{hard}|^2$) consisting of a 4 momentum change from p' to p through a gluon emission with energy (E_k) at

angle of θ. The cross-section of such a process is the product of the amplitude and the phase space (Φ):

$$|M|^2 d\Phi \approx |M_{hard}|^2 d\Phi_{hard} \frac{2C_F \alpha_S}{\pi} \frac{dE_k}{E_K} \frac{d\theta}{\theta} \tag{1.7}$$

which factorizes into hard matrix element and a soft gluon emission probability:

$$P_{\text{soft-gluon-emission}} = \frac{2C_F \alpha_S}{\pi} \frac{dE_k}{E_K} \frac{d\theta}{\theta} \tag{1.8}$$

where C_F is the color factor representing the probability of the coupling of the gluon to a quark ($C_F = 4/3$) while replacing it with the color factor $C_A = 3$ instead can represent the coupling of a gluon to another gluon. There are divergences for this probability for soft ($E \to 0$) and collinear ($\theta \to 0$) limits. This leads to a large emission cascade of soft collinear partons, which are primarily gluons due to the color factors. One can indeed check that the average number of emitted gluons, obtained by integrating the soft probability of the previous equation, is large:

$$\langle N_{\text{gluons}} \rangle \approx \int P_{\text{soft-gluon-emission}} \approx \frac{1}{\alpha_S} \gg 1 \tag{1.9}$$

The parton shower is the process of quarks and gluons consecutively radiating partons. The formation of this cascade of partons is driven by the infrared and collinear divergences of QCD and the running of the strong coupling.

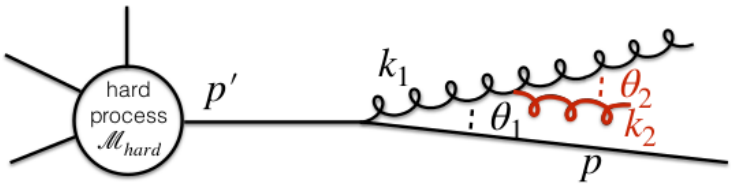


Figure 1.9: Schematic of angular ordered gluon emissions showing a second gluon emission. Figure extracted from Gavin Salam's lecture notes on QCD [10]

In Fig. 1.9, a sketch for the process consisting of a second gluon emission is shown. The

emission of a second gluon factorizes from the emission of the first. However coherence effects manifest in angular ordering: the second gluon is confined to $\theta_2 < \theta_1$

The cross section for the emission of an extra gluon can be rewritten in terms of the radiation function W_{ij} that can be shown to follow strict angular ordering:

$$d\sigma_{n+1} = d\sigma_n \frac{d\omega}{\omega} \frac{d\Omega}{2\pi} \frac{\alpha_S}{2\pi} \sum_{i,j} C_{ij} W_{ij} \tag{1.10}$$

The radiation function given by W_{ij} is,

$$W_{ij} = \frac{\omega^2 p_i \cdot p_j}{p_i \cdot q \, p_j \cdot q} = \frac{1 - v_i v_j \cos\theta_{ij}}{(1 - v_i \cos\theta_{iq})(1 - v_j \cos\theta_{jq})}. \tag{1.11}$$

This cascade of gluons emitting further gluons is thus confined within a cone $\theta_{i+1} \lesssim \theta_i$; the emissions are angular ordered as shown in Fig. 1.9. Angular ordering accounts for the appropriate dependence of the soft gluon radiation on the prehistory of parton shower development [11].

The parton cascade progresses until the scale of the emissions reaches nonperturbative values (or equivalently, when the transverse separation between partons becomes of the order of the size of the hadron). Then, hadronization happens, as depicted in Fig. 1.10. Hadronization cannot be described analytically with perturbative tools and encodes the transition from the fundamental degrees of freedom of QCD to color-neutral composite particles in the final-state, i.e. the hadrons and their decays that can be directly measured experimentally. There exist two major model descriptions of this nonperturbative hadronization process which connects the experimentally inaccessible partons to the final state hadrons:

- *Lund string fragmentation model* [12]: The Lund string model is based on the classical interaction potential V between partons, which can be described

by a potential that depends linearly on their separation r based on lattice QCD calculations. The interactions between partons can be analogous to a string connecting them. An increase in the separation between the partons corresponds to an increase in the attractive potential stored in the string. It is used in MC models such as PYTHIA[1].

- *Cluster fragmentation model* [13]: The cluster fragmentation model is based on the color confinement property of QCD, where partons are clustered into color singlet groups. The groups are then involved in the hadronization process with a given invariant mass and decay to other color singlet states. It is used in MC models such as SHERPA and HERWIG[1].

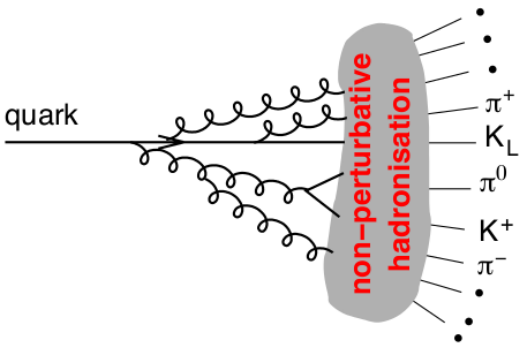


Figure 1.10: Schematic of the nonperturbative hadronization process showing the transition from a parent quark to collimated sprays of color neutral particles called "jets". Figure extracted from Gavin Salam's lecture notes on QCD [10]

The parton shower and hadronization process leads to the observation of multiple collimated hadrons originating from the initial energetic parton which are called jets. These jets contain the initial state information of the parton and are one of the best probes of fundamental principles of QCD. Other common methods of quantitatively studying QCD are through heavy quark production, measurement of the running coupling and event shape observables. Among these methods, the focus of this thesis is the study of

[1]PYTHIA8 and HERWIG7 are the MC models used in this thesis. Further description about their usage is provided in Chapter 4

jets and their internal structure to provide information about the parent parton.

1.2.5 Definition of a jet

One way of studying QCD at short distance scales is via the study of the production of highly collimated sprays of particles known as "jets". Jets are collimated flows of hadrons and they can be seen as proxies to the high-energy quarks and gluons produced in a collision. Since this does not involve knowledge about the initial parton, the definition of a jet needs to be resilient to QCD effects. A definition of a jet consists of two main ingredients: a jet algorithm and a recombination scheme. In hadronic collisions with a multitude of particles, the question of which particles to put together in to a given jet and the method to combine their momenta is specified by the jet clustering algorithm and definition. The recombination scheme specifies how the kinematic properties of the constituent particles are combined to give the properties of a complete jet. Some of the common methods used are *E-scheme* which considers the sum of the 4-momenta of the individual particles and is the most widely used combination scheme, *winner-take-all* (WTA) where the resulting jet is along the axis of the largest p_{T} constituent with the sum of all constituent p_{T}. Other methods for recombination include the massless p_{T} and E_T schemes.

The final state particles which are clustered to form jets are not limited to the hadronization components of a hard scattering. Hadrons formed from detector noise, beam-beam remnants, multiple parton interactions as well initial and final state radiation are collectively known as the underlying event (UE) which forms a uniform background for each event. In addition, due to the collision involving a bunch of particles in the beam, there can be interactions other than the primary hard scattering known as pileup. Jets are affected in various ways by these processes depending on their algorithm and a resilience to these background effects can be a useful tool in the study of QCD.

Any combination of particles could be considered a jet, but a useful definition of a jet produces the same result in the face of the divergent limits of the QCD to allow perturbative calculations. For an observable's distribution to be calculable in fixed-order perturbation theory, the observables should be infrared safe, i.e. insensitive to the emission of soft or collinear gluons. In particular, $\vec{p}_i$ is any momentum such that in its definition, it must be invariant under the branching,

$$\vec{p}_i \rightarrow \vec{p}_j + \vec{p}_k \tag{1.12}$$

whenever $\vec{p}_j$ and $\vec{p}_k$ are parallel (collinear) or one of them is small (infrared) [14]. Consequently, for a jet algorithm to be well defined in pQCD, it needs to be infrared and collinear (IRC)-safe.

Earliest definitions of jets often employed *cone algorithms* which required the constituents to be within a cone of predefined radius. While these algorithms produce jets which trace out rigid circles, earlier algorithms such as the Iterative Cone with Split-Merge procedure (IC-SM) [15] and Iterative Cone with progressive removal(IC-PR) [16] were infrared and collinear unsafe, respectively. The Seedless Infrared Safe Cone (SISCone) algorithm [17] was developed to be an IRC safe cone algorithm.

The other class of jet clustering algorithms are known as *sequential algorithms*[18–20]. Sequential recombination algorithms utilize both the momentum and the angular separation between particles when combining them into a jet. The k_{T}, anti-k_{T} and Cambridge–Aachen (CA) algorithms for jet clustering are among the most commonly used sequential recombination algorithms.

These algorithms rely on calculating two distance measures, $d_{i,j}$ and $d_{i,B}$ for every par-

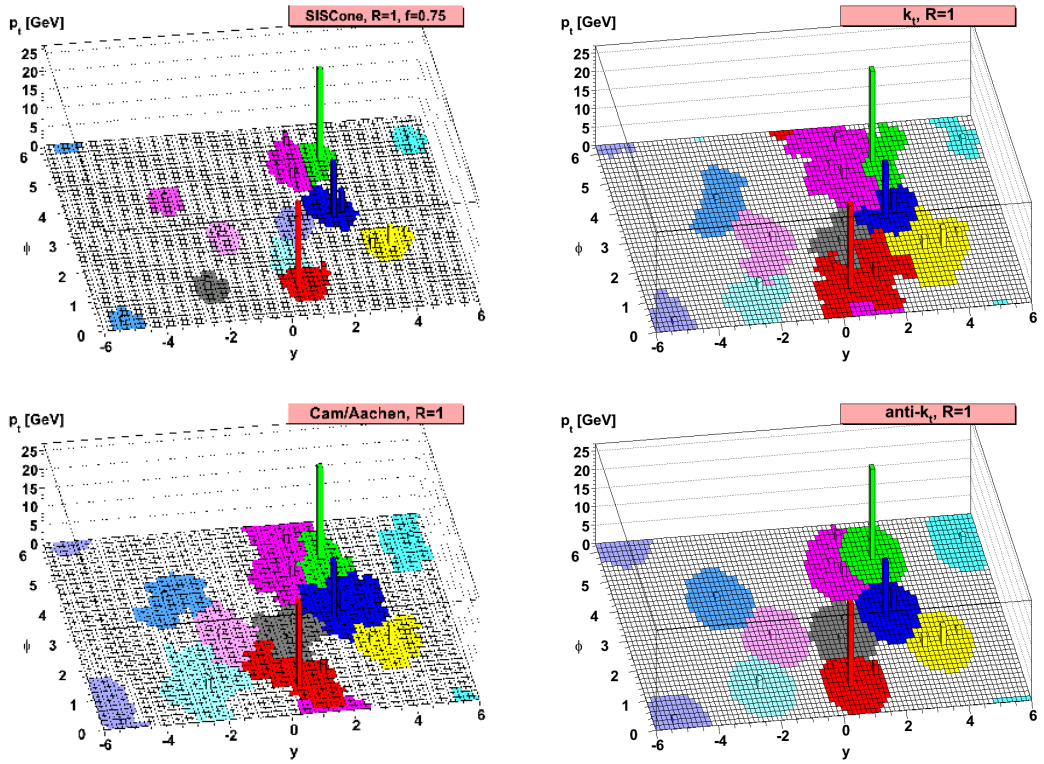


Figure 1.11: Multiple IRC safe jet clustering algorithms are illustrated using a sample parton level event along with multiple soft emissions. The algorithms shown are SISCone(upper left), k_{T} (upper right), anti-k_{T} (lower right) and CA (lower left). The characteristics of each algorithm are seen in the active areas for each hard jet. Notably the SISCone and anti-k_{T} algorithms produce circular jets unlike k_{T} or CA whose structure is dependent on the soft particles [18]

ticle defined as,

$$d_{i,j} = min(p^a_{T,i}, p^a_{T,j})\frac{R^2_{i,j}}{R^2} \qquad d_{i,B} = p^a_{T,i} \tag{1.13}$$

where i and j are the particles, $R_{i,j}$ is the Lorentz boost invariant distance and p_T is the momentum transverse to the beam axis. The $d_{i,B}$ is the momentum space distance between the ith particle and the beam axis. $d_{i,B}$ is the limit such that when $d_{i,j} < d_{i,B}$, the ith and jth particle are combined into one. Further details about the coordinate system used in the definition of these variables are provided in Chapter 2. The parameter a is

used to distinguish different classes of sequential clustering algorithms as follows and shown in Fig. 1.11,

The k_T algorithm: The case where $a = 1$ is the k_T algorithm. The k_T algorithm begins with the clustering of soft particles and includes the hard component in the final steps of the clustering as it sorts in increasing order of p_{T} of particles within the jet cone radius R. The k_T algorithm inverts the QCD branching process as the distance parameter $d_{i,j} \approx \frac{1}{P_{ij}}$ where P_{ij} is the emission probability in QCD. This leads to the k_T algorithm producing irregularly shaped jets.

The anti-k_T algorithm: The case where $a = -1$ is the anti-k_T algorithm. It is the inverse of the k_T algorithm by sorting the p_{T} of particles starting from the highest to the lowest. Thus, the anti-k_T algorithm clusters the hardest particles first, leaving the soft particles to be clustered at the final steps resulting in approximately circular shaped jets. This feature is very useful in the removal of pileup, for jet calibrations and UE subtraction, similar to cone algorithms thus being highly favored in experimental and theoretical communities.

The Cambridge–Aachen (CA) algorithm: The case where $a = 0$ is the CA algorithm. By removing the p_{T} dependence on the jet clustering, the CA algorithm is uniquely angular ordered. The clustering is purely geometrical, depending only on the distribution of particles, starting from the closest to the furthest. Thus, it is useful when studying jet substructure as discussed in the following section.

Looking inside jets: jet substructure

The study of the internal structure of jets using substructure techniques can been employed in experimental analyses. The tools developed to study jet substructure have applications for SM measurements as well as search for new physics. These techniques have renewed the efforts in the collaboration of experimental and theoretical studies

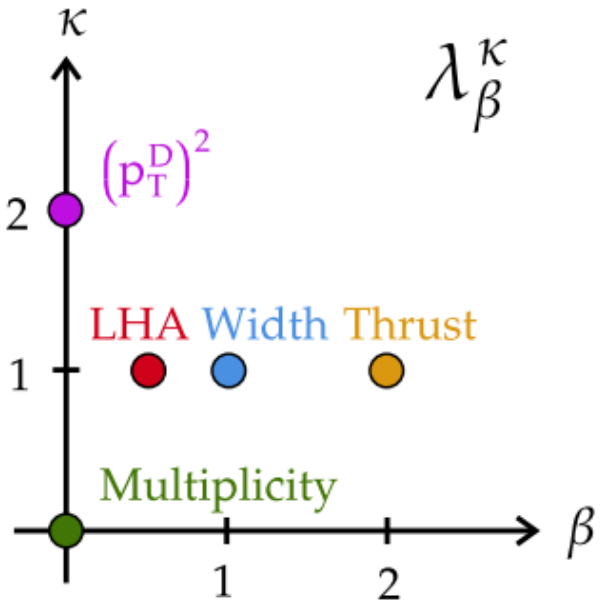


Figure 1.12: Schematic of the various angularities by varying the parameters β and κ in equation 1.17 [21]

by designing better observables to describe the QCD at higher orders. One of the most common use for jet substructure tools is to disentangle various kinds of jets such as, quark and gluon initiated jets or isolating boosted massive bosons. There are two broad classification of jet substructure observables [22]:

- Observables built using the position and momenta of the jet constituents such as jet fragmentation functions [23–25] (eq. 1.14), the $N-$subjettiness [26] (eq. 1.15) or energy-energy correlators [27, 28] (eq. 1.16) or generalized angularities (shown in eq. 1.17 and fig. 1.12). Examples of measured angularities [29] are the jet mass, the jet girth (g), and the momentum dispersion $p_{T,D}$.

$$z \equiv \frac{p_T^{\text{part}} \cos \Delta R}{p_T^{\text{jet}}} \tag{1.14}$$

$$\tau_N^{(\beta)} = \sum_{i \in \text{jet}} p_{ti} \min(\Delta R_{ia_1}^{\beta}, \cdots, \Delta R_{ia_N}^{\beta}) \tag{1.15}$$

$$e_N^{(\beta)} = \sum_{i_1<\cdots<i_N\in\text{jet}} \left(\prod_{j=1}^{N} z_{i_j}\right) \left(\prod_{k<\ell=1}^{N} \Delta R_{i_k i_\ell}^{\beta}\right) \tag{1.16}$$

$$\lambda_\beta^\kappa = \sum_{i\in\text{jet}} z_i^\kappa (R_i/R_0)^\beta \tag{1.17}$$

- Another class of observables is built using the clustering history of the jets. The history of the jet clustering, depending on the jet algorithm provides access to the information about the energy flow via distinguishing soft and hard radiation components of the jet. These are accessed through observables such as groomed momentum balance z_g, the groomed jet radius (R_g) or the Les Houches multiplicity (n_{SD}).

The substructure observables constructed using the clustering history can employ iterative declustering techniques to study the evolution of the jet shower. The constituents of the jet are usually reclustered using an angular ordered jet clustering algorithm known as the CA algorithm described in the previous section. The CA jet obtained is then iteratively declustered following the hardest subject, registering the kinematics of the subleading branch at each step in order to build the Lund jet plane. The Lund plane is the most general observable built with the jet tree providing a representation of the phase space of $1 \to 2$ branchings and can be constructed through the iterative declustering procedure. It is constructed to represent the density of emissions as a function of the splitting angle (ΔR) and the transverse momentum of the emission (k_T). Figure 1.13 [30] shows a schematic of the different regions explored by the Lund plane and a recent measurement of it in proton-proton collisions with CMS.

The iterative declustering can be complete as shown in the Lund plane or it can be partial until a node that satisfies a given kinematic condition is reached. This procedure of premptively stopping the declustering to probe various regions of the the Lund plane is known as "grooming". An example of grooming is the Soft Drop (SD) grooming

algorithm with the condition on the momentum fraction z, resulting in a groomed jet as shown in Fig.1.14,

$$z > z_{cut}(R/R_0)^{\beta} \tag{1.18}$$

where $z = \frac{j_{2,p_T}}{j_{1,p_T}+j_{2,p_T}}$ such that j_1 and j_2 are the leading and sub-leading subjets at each step of the jet declustering. The grooming procedure is representative of a diagonal cut

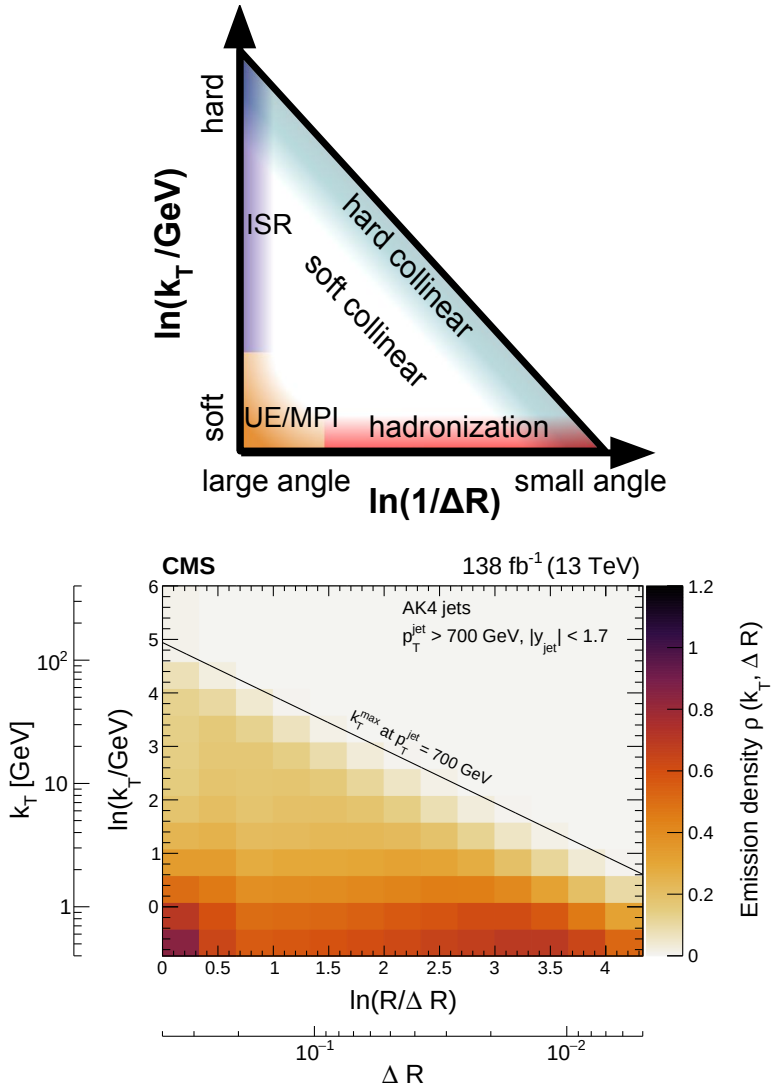


Figure 1.13: Schematic diagram of the Lund jet plane(upper) affected by the different mechanisms populating the various regions and a measurement of the Lund jet plane in pp collisions by CMS (lower) is shown [30]

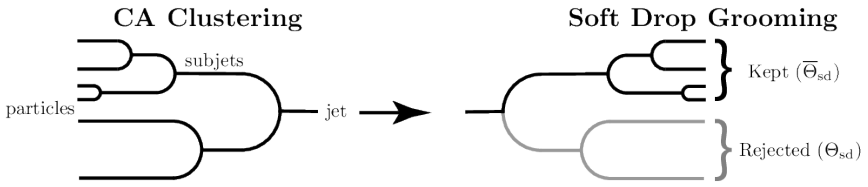


Figure 1.14: Schematic diagram of soft drop grooming a Cambridge/Aachen jet [31]

in the Lund jet plane. The free parameters z_{cut} and β are chosen such that higher z_{cut} values improves the purity of the measured prongs. Applying the SD condition, allows the "groomed" observables to not be affected by soft emissions and UE effects. Other grooming approaches include Mass Drop [32], Recursive Soft Drop [33] and Dynamical k_T [34].

The results presented in this thesis utilize the jet girth g and the groomed jet radius (R_g) which are representative of the two classes of substructure variables and are described in greater detail in Section 4.5.

1.3 Quark Gluon Plasma

In the previous section, we have discussed multiple aspects of QCD and properties of quarks and gluons. An in depth look into the behavior of QCD matter is possible by looking at the quark-gluon plasma (QGP). The QGP is a deconfined state of matter where the quarks and gluons are in a localized approximate thermal and chemical equilibrium. The presence of free color charges are allowed making this state of matter a plasma. Due to the self interaction of the gluons, the running of the strong coupling and confinement, there is a wide range of emergent phenomena in the strong interaction sector such as the origin of the hadron mass scale, chiral symmetry breaking and the pattern of hadronic bound states. Understanding these emergent phenomena is one of the long-term goals in high energy physics and describing the vast phase space with the fundamental theory of QCD is non-trivial. Probing the QGP is various regimes of temperature and net baryon density can be used to improve our understanding of the fundamental principles of QCD. The QGP is the state of matter which was present in the early universe at a timescale of the first 10^{-6}s after the Big Bang, when the density and temperatures allowed the presence of deconfined quarks and gluons before the expansion of the universe cooled to form stable hadrons. Direct observable studies of the early universe are performed on the Cosmic Microwave Background which is the state of the equilibrium after the QGP hadronizes. Thus, direct studies of the evolution of the QGP during the early universe is not possible. These extreme conditions required are naturally present in the core neutron stars and in gamma ray bursts. A laboratory recreation of the conditions present in the early universe is possible in relativistic heavy ion collisions used to study nuclear matter. In this section, we discuss the connection between the QGP, its properties and how it is used in studying the fundamental principles of QCD.

1.3.1 QCD phase diagram

Changes in degrees of freedom of matter is usually described by a phase diagram which is a graphical representation of the state of a substance under different temperature and pressure conditions. The behavior of quarks and gluons at various conditions is represented by the QCD phase diagram shown in Fig. 1.15. The axis of the phase diagram are the net baryon density and the temperature. There are multiple regions formed in the QCD phase diagram with the experiments probing each scale shown in the figure. Stable matter is present as hadrons and atomic nuclei in the bottom left at low temperature and density. We approach the astrophysical regime of high net baryon densities seen in neutron stars. Of particular interest in this thesis is the transition at low baryon densities and high temperature. A crossover transition exists where stable nuclei converts into a deconfined phase of quarks and gluons; the QGP. A complete description of the QCD phase diagram requires a thorough understanding of dense, strongly interacting hadronic matter which is in the nonperturbative regime of QCD and is theoretically challenging to describe.

A few theoretical approaches used in the understanding of the QCD phase diagram are weakly coupled theory, macroscopic models such as the bag model, effective field theories or lattice gauge theory. Lattice gauge theory or lattice QCD in particular is a first principles calculation tool describing the thermodynamic properties of the QGP using discretized lattice points in space and time, and directly solving the equations of motion of QCD numerically. By calculating the QCD equation of state in equilibrium for a series of lattices, the theory approaches a description of QCD in the continuum limit for small lattice sizes. This is performed using Monte Carlo (MC) simulations and is a computationally intensive process generally applicable at low densities. Predictions from lattice QCD have been successfully validated through multiple experimental results such as determining the mass of the proton [36], and other hadrons along with the

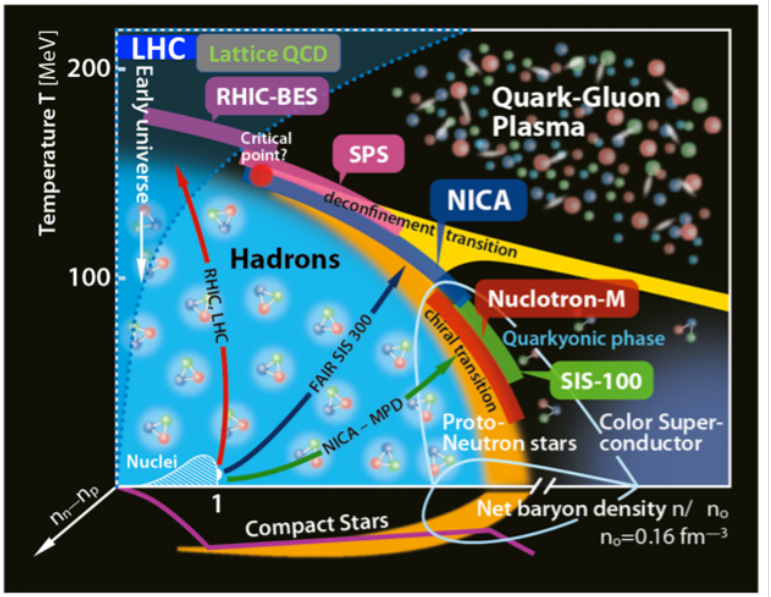


Figure 1.15: This figure shows the QCD phase diagram with the horizontal axis consisting of the net baryon density and the vertical axis, the temperature. A phase diagram describes the different states of matter as shown such as stable matter in the form of hadrons and nuclei in the low temperature and density regime, neutron stars and color superconductors in the low temperature high density regime and a transition to a deconfined state, the QGP in higher temperatures or density regimes. Various experiments and the regimes probed by them are shown [35]

fact that the transition from stable hadrons to a deconfined QGP is a crossover and not a first order phase transition around a temperature of 150 MeV [37]. It was predicted in the 1960s by Hagedorn [38] in a effective theory to explain the number of resonance states diverged at this temperature which predates the advent of QCD.

An alternative theoretical conjecture formulated using string theory and quantum gravity known as the anti-de Sitter/conformal field theory correspondence (AdS/CFT) provides a powerful toolkit for studying the strongly coupled quantum field theories. It provides the best understood nonperturbative formulation of string theory with specified boundary conditions. This formalism is extremely useful in describing the QGP which is known as the perfect fluid. This is due to measurements of the specific viscosity of the QGP are compatible with $1/4\pi$ predicted by AdS/CFT and is smaller than

any other known fluid [39]. The value $1/4\pi$ is also the specific viscosity of plasma in strongly coupled gauge theories.

A complete understanding of the QCD phase diagram with these theoretical frameworks would be required to answer the open questions posed by the experimental observations in the various regimes QCD: how does a strongly coupled liquid emerge in the asymptotically free limit of quantum gauge theories? What is the limit of size and timescale when the QGP can be described by hydrodynamics? How does the transition between a hydrodynamic liquid to a non-hydrodynamic happen? A discussion about the studies performed to uncover some of these is presented in the further sections, primarily in the context of the physics of relativistic heavy ion collisions probing the low density and high temperature regime of the QCD phase diagram. However, a brief description on a complementary probe of the high density regime of the phase diagram available in nature in the form of neutron stars is first presented before discussing the laboratory environment of relativistic collisions to provide context.

1.3.2 Astrophysical implications

As shown in the QCD phase diagram in Fig. 1.16 [40] neutron stars can be used to probe the high baryon density regime of the QCD phase diagram. The internal structure of neutron stars is not yet well understood. Descriptions of the equation of state of neutron star mergers can lead to temperature scales that overlap with the lowest energy heavy ion collisions providing a comparison to extend our understanding of the phase diagram. Neutron star mergers are events in which all four fundamental forces are combined.

The study of neutron star mergers is performed with simulations of general relativistic hydrodynamics. These events are a primary source of gravitational waves and gamma ray bursts. Several constraints can be placed on the nuclear equation of state in re-

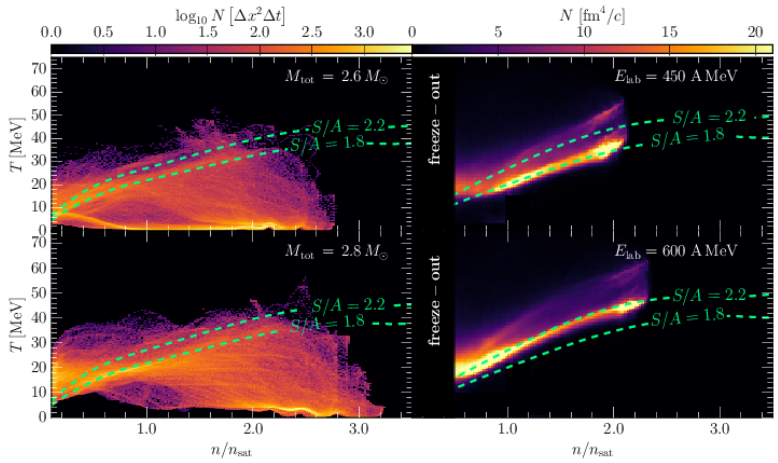


Figure 1.16: Simulation studies showing regions of the QCD phase-diagram probed by two binary neutron star mergers with different masses (left) and by two heavy ion collisions with different beam energies (right) [40]

cent years from neutron star observations, laboratory experiments and theoretical models [41–45].

1.3.3 QCD in Heavy Ion Collisions

The QCD matter has multiple thermodynamic phases such as a stable hadrons, as a high density color superconductor, a transition quarkyonic phase causing deconfinement and the quark gluon plasma as shown in Fig.1.15. Thus, the question about the presence of a critical point is a fundamental question about the properties of the QGP. It has been shown through the first principles lattice QCD calculations that there is a continuity between the two phases of hadron gas and the QGP at zero baryon density [37]. The identification of the critical point through lattice calculations will be a milestone in the theoretical framework of QCD calculations. It is also one of the questions to be addressed experimentally by scanning the phase diagram at various collisions energies as shown in Fig.1.17 [46].

There are two available quantities to control in heavy ion collisions: the species involved in the collision and their energy. The energy of the collision can be known to a

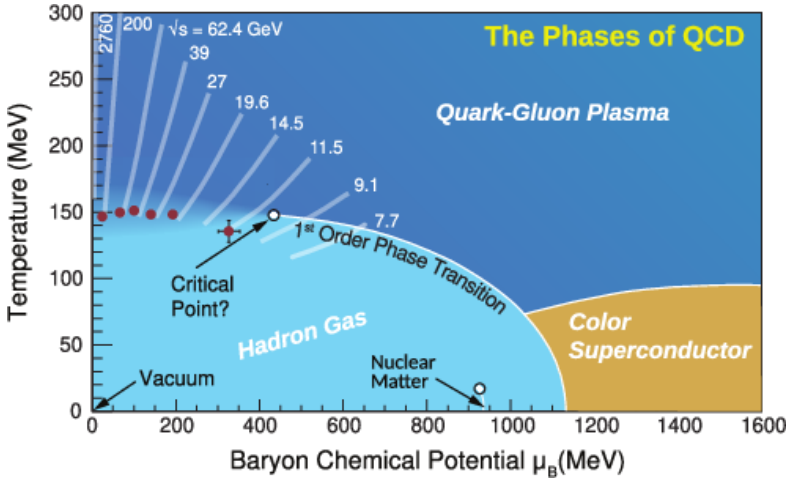


Figure 1.17: QCD phase diagram showing the regions probed by beam energy scans at RHIC [46]

very high precision but in nuclear collisions the geometry and composition of the collision itself can play a part in the physics processes studied in heavy ion collisions. The geometry of the collision is characterized by the impact parameter b, which is the transverse distance between the center of masses of the colliding nuclei. An approximation for the distribution of quarks and gluons in the nuclei is also made by considering the PDFs for nuclear collisions to be the same as for free nucleons as described by $f_{a,b}(x)$ in equation 1.6 for the hadronic production cross section [47]. In addition it is possible to model the nucleons as hard spheres with an energy dependent radius by considering the total inelastic pp cross section [48] as shown in the Glauber Model in Fig.1.18 [49]. The Glauber model calculations have two main classes: the "optical limit" calculations where a smooth matter density is assumed inside the nuclei, and the "Glauber Monte Carlo" where the nucleus is modelled as uncorrelated nucleons sampled from measured density distributions. The thickness function of the nuclei involved in the collision can be expressed as,

$$T_A(s) = \int dz \rho_A(s,z) \tag{1.19}$$

where, ρ is the nuclear density distribution for a given nucleus with A nucleons normalized such that $\int d^2sT_A(s) = A$. This is parameterized using the Wood-Saxon distribution to give the nuclear overlap function,

$$T_{AB}(b) = \int d^2sT_A(s)T_B(s-b) \tag{1.20}$$

Further description for the experimental determination of the collision geometry defined by centrality is discussed in Section 4.2.

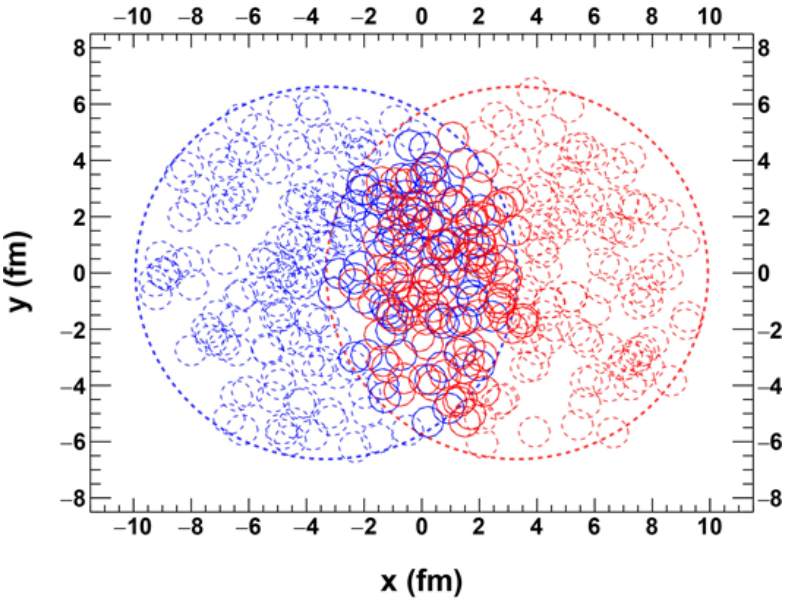


Figure 1.18: Simulation of a PbPb collision for the Glauber model with impact parameter $b \approx 7$ fm [49]

The QGP formed in heavy ion collisions can be studied through two kinds of experimental probes:

Soft probes: Soft processes involve low momentum transfer depending on the parton densities of the colliding nucleons. Most of the particles produced in heavy ion collisions are the result of soft collisions. These soft particles are a consequence of the global properties of the evolution of the QGP and allow for the hydrodynamic modeling of the

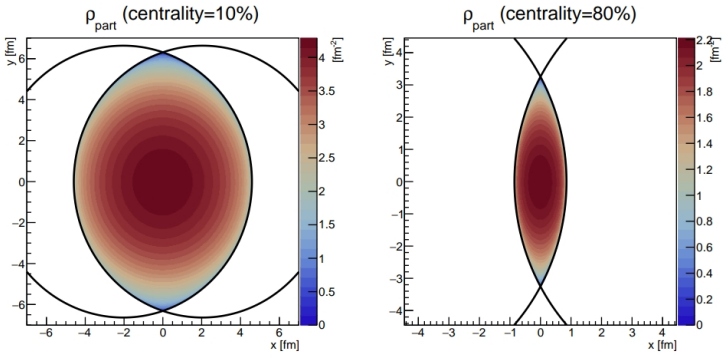


Figure 1.19: The overlap region in PbPb collisions for two different centralities. The shape of the overlap is an oval which leads to anisotropies in the evolution of the QGP. [50]

QGP. The hydrodynamical evolution of the medium results in highly correlated final-state particle distributions that result in long range correlations. The colliding participants in heavy ion collisions takes an oval shape as shown in Fig. 1.19 causing an azimuthal anisotropy which is observable in the coefficients of Fourier decomposition of the particles. The large pressure gradients in the transverse direction to the event collision drives the expansion of the medium causing the elliptic flow (v_2) to become dominant. In addition to elliptic flow, higher order Fourier coefficients such as v_3 and v_4 represent triangular flow and quadrangular flow and are sensitive to the initial stage of the evolution. Thus, they are important in the study of hydrodynamical properties of the QGP such as viscosity and temperature. One of the recent measurements shown in Fig. 1.20 is of the speed of sound in the strongly interacting QGP [51]. This measurement relied on the analysis of ultra-central event caused by head-on heavy ion collisions and the study of the spectra of very low p_{T} particles produced in these collisions.

Hard probes: Processes that involve high p_{T} particles produced in the early stages of the collision such as heavy quarks and boosted electroweak bosons constitute hard probes. These particles experience the full evolution of the QGP while being perturbatively calculable due to the large momentum transfers involved in their production. There are generally two kinds of hard probes based on the presence of color charge and conse-

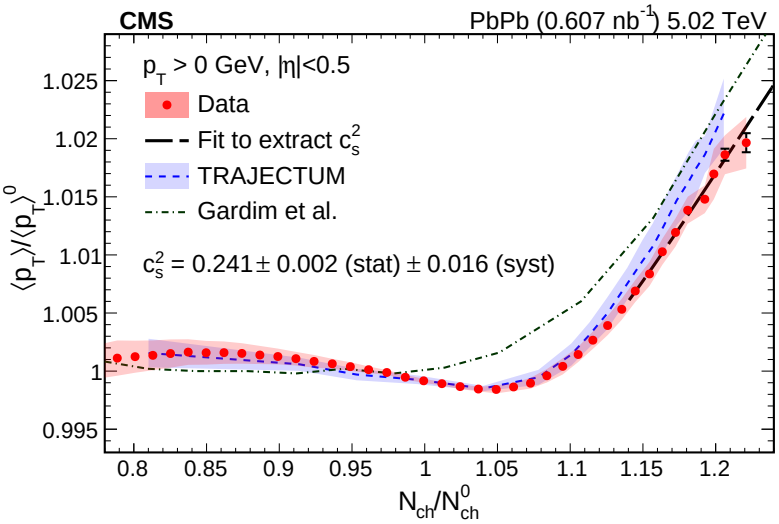


Figure 1.20: The average transverse momentum of charged particles as a function of the charged-particle multiplicity. The dashed line is the fit to data used to extract the speed of sound c_s [51]

quentially the ability to interact with the QGP through the strong force. Quarks and gluons which carry color charge are colored probes that can provide direct information about the partonic interaction with the QGP. Colorless probes such as the photons, W bosons and Z bosons on the other hand, only interact through the electroweak force and can be used as a control for the initial state of the collision in comparisons to the pp reference.

Jets in heavy ion collisions

The evolution of the vacuum parton shower shown in Fig. 1.7 is modified in relativistic heavy ion collisions. Due to formation time arguments, hard nucleon-nucleon interactions leading to highly virtual parton formation happen early in time as compared to other soft interactions leading to the QGP formation. As a consequence, the elementary scattering and the initial state evolution encoded in the PDF's in eq. 1.6 describing jet production, remain vacuum or pp-like in heavy ion collisions. As the energetic parton shower develops, it will encounter the time-evolving QGP medium and their evolution will be intertwined. Jets provide a "tomographic view" of the QGP medium.

The fragmentation of highly energetic jets, initiated by high-virtuality partons produced in the early stages of the collision, is intertwined with the time evolution of this hot and dense matter. The interactions between the jet and this medium happen over a wide range of energy scales, from the high virtuality of the initial parton of up to hundred GeV to the medium temperature of hundreds of MeV. These interactions are responsible for partonic energy loss in the QGP medium, a phenomenon generally termed as "jet quenching." These interactions are also expected to modify the jet radiation pattern in PbPb collisions when compared with pp collisions.

The interactions between the jet and the QGP medium happen over a wide range of energy scales, from the high virtuality of the initial parton of up to hundred GeV to the medium temperature of hundreds of MeV. They do not only involve the modification of the radiation pattern of the individual partons but also the interactions between several emissions within the medium and the response of the medium induced by the jet.

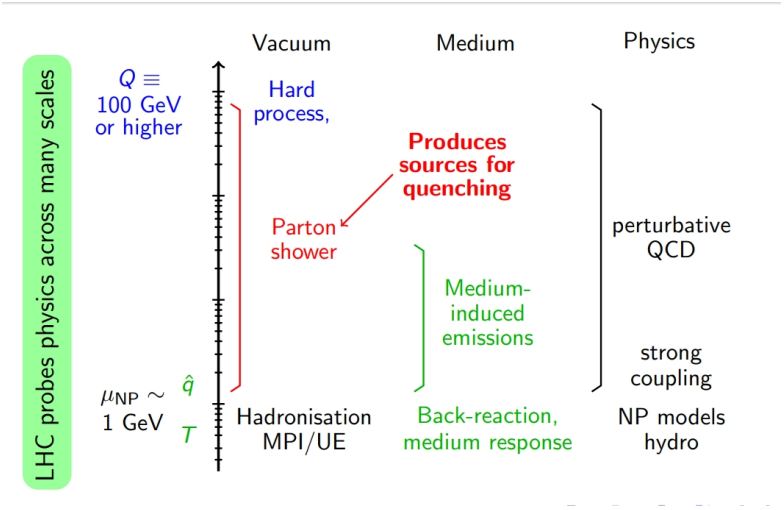


Figure 1.21: Probing the physics of QCD across many scales at the LHC. Adapted from Gregory Soyez's slides [52]

Figure. 1.21 describes the various processes involved in the parton-QGP interactions in the different energy scales.

A hard elementary interaction produces a high-Q parton. This parton will reduce its virtuality by radiating in vacuum (DGLAP evolution). At some point, this early vacuum shower will encounter the QGP medium and its constituents will act as sources for quenching. The main mechanism for quenching at perturbative scales is medium-induced radiation of soft gluons. As the medium-modified shower evolves, the shower constituents get softer, and their subsequent interactions with the medium occur at lower scales, eventually of the same order as the temperature of the medium. Perturbative techniques cannot be applied at such low scales and the physics of this domain can be described by strong coupling calculations based on the AdS/CFT conjecture as the Hybrid model[53]. Finally, hadronisation time scales are long and of the order of $\frac{1}{\Lambda_{QCD}}$ and is usually considered to be unaffected by the medium.

The picture above considers a factorization of scales according to which the vacuum shower happens earlier than the medium-modified shower. This picture is valid in a leading, double-logarithmic approximation as was shown for the first time in [54].

As mentioned above, energetic partons propagating through the QGP are expected to lose energy mainly via radiative processes[55–57]. These processes radiate gluons due to the scattering of energetic partons with the QGP medium. The spectrum of the single gluon emission can be calculated from first principles using in-medium propagators describing the scattering of the projectile parton with the static scattering centers of the QGP. Two approximate limits can be considered to allow for a semi-analytical treatment: the limit where the interactions with the medium are few hard scatterings (referred to as the opacity expansion [58],[59]) and the limit where the interactions are multiple soft interactions with the medium (referred to as the harmonic oscillator approximation [56, 57, 60]).

In this latter case, the amount of energy loss is controlled by a single parameter, the transport coefficient $\hat{q}$, which represents the average momentum given to the parton by

the medium per unit path length, with units of $\mathrm{GeV}^2/\mathrm{fm}$ given by,

$$\hat{q} = \frac{\langle k_T^2 \rangle}{\lambda} \tag{1.21}$$

where λ is the mean free path of the QGP medium. The medium-induced emission kernel shows a characteristic gluon emission frequency $\omega_C = \hat{q}L^2/2$; where L is the length of the medium. It shows Landau-Pomeranchuck-Migdal coherent suppression for emission formation times that are longer than the medium length [61].

An example of recent theoretical effort to overcome the limitations of the two approximations is the Improved Opacity Expansion [62, 63], which includes both single hard and multiple soft scatterings. See also [64] numerical developments. A medium-induced gluon spectrum that incorporates both the soft and hard limits are essential as a theoretical reference for probing the shortest length scales in the QGP, including answering whether it is possible to resolve point-like scattering within the QGP.

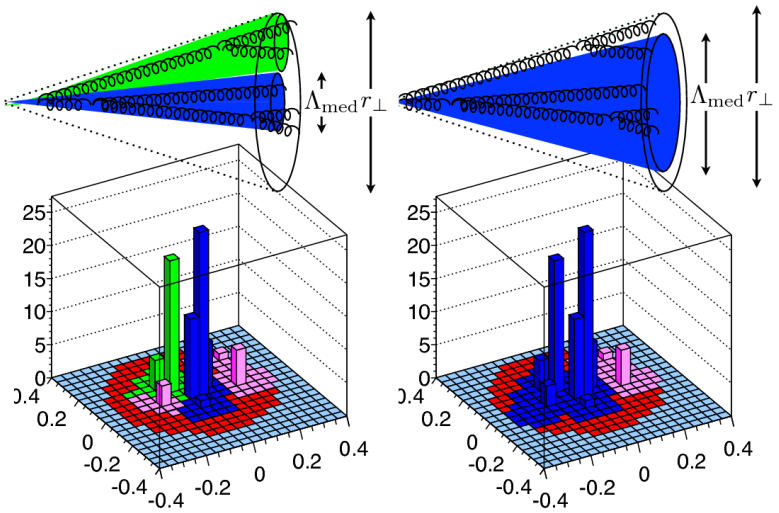


Figure 1.22: A schematic diagram of a jet with its constituents representing the medium resolution length. The blue histogram denotes the hardest resolved sub-jet, the green the next-to-hardest one, while the pink histogram denotes soft fragments. [65]

To assess color coherence in multigluon emissions in medium, the medium-induced

gluon off a $q\bar{q}$ antenna was calculated. These calculations exposed a new transverse scale, the "medium resolution length" [66–68]. The medium resolution length is the minimum angular separation between two partons or subjects that can be resolved by the medium as independent color charges. This concept is schematically shown in Fig. 1.22 [65]. The existence of this new scale has a clear phenomenological consequence: jet energy loss will depend on the internal structure of jets. Jets with larger number of resolved constituents will be more quenched.

The medium resolution length is defined as,

$$\Lambda_{\text{med}} = 1 - \exp\left(-\frac{1}{12}\hat{q}\theta_{q\bar{q}}L^3\right) \tag{1.22}$$

where $\hat{q}$ is the quenching parameter described previously, L is the length of the medium, and $\theta_{q\bar{q}}$ is the opening angle of the antenna. In a dense medium, $\Delta_{med} \to 1$ and angular ordering of the emissions is broken.

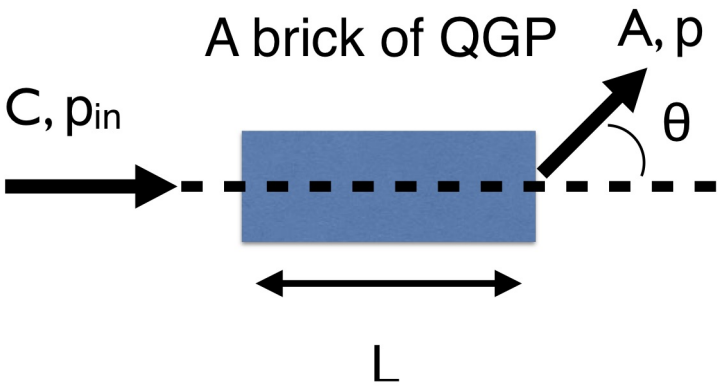


Figure 1.23: A schematic diagram of a hard parton scattering in the QGP. An incident parton C interacts with the QGP of length L resulting in an outgoing parton A at an angle θ. The momentum is changed from p_{in} to p [69]

An underlying assumption in most of the theoretical developments described above is that the QGP scattering centers are static. When a parton interacts with the QGP scattering centers, those are expected to be deflected and can end up as part of the

jet. One can view this as the medium that gets excited by the jet passage, referred to as back-reaction or medium response in the Fig. 1.21. This contribution is soft, and subject to modeling.

The developments mentioned earlier rely on perturbative QCD and are not applicable in the scales where the QGP temperature for a weak coupling is not valid. This is resolved in the Hybrid Model [53] by treating physics processes at different energy scales differently. Since we'll perform extensive comparisons of our results to the Hybrid model later in this manuscript, we take the time now to briefly describe its main ingredients. The perturbative process of the single parton fragmentation up to hadronization are treated following the DGLAP evolution [70–72], in practice using PYTHIA. The soft interactions with the medium are then modeled considering energy loss of quarks and gluons embedded in a strongly coupled hydrodynamic evolution. The Hybrid model includes the modeling of energy loss due to the strong coupling between the partons and the medium, described as a drag each parton constituent experiences as it propagates through the medium, and the large-angle deflections of partons traversing the QGP, referred to as "Moliére elastic scatterings" [69, 73]. The nonperturbative backreaction of the medium [74], also known as medium response, is also included in the model. The Moliére scatterings are denoted as "elastic" and the medium response as "wake" and discussed further in Chapter 6 in the context of this thesis. These elastic scatterings can be thought to be analogous to a scattering experiment as shown by Fig. 1.23 [69]. The key point is that a scattering of a projectile parton with a strongly coupled medium can only lead to a gaussian distribution of deflected angles or momentum transferrd k_T. A scattering with a large momentum transfer with the medium must involve point-like scattering centers. In other words, a hard scattering in the medium is a signature of a scattering against the free quarks and gluons within the strongly coupled QGP. The model also considers an ad hoc implementation of color coherence called L_{res} which is discussed in Chapter 6.

These interactions taken holistically are expected to modify the jet radiation pattern in nucleon-nucleon collisions when compared with pp collisions. Jets produced in pp collisions, which are unaffected by the medium, are used as a reference for "vacuum radiation." Thus experiments rely on comparisons of observables in the 2 collision systems to study the properties of the QGP.

1.3.4 Experimental probes for jet quenching

Inclusive jet and jet correlations:

One of the observables used to quantify differences in heavy ion collisions with respect to the vacuum reference of pp collisions is the nuclear modification factor. It is the ratio of yields in nucleon-nucleon collisions to the corresponding yield in pp collisions scaled by the number of binary collisions $\langle N_{\text{coll}} \rangle$ defined as,

$$R_{AA}(p_T) = \frac{\sigma_{NN}}{\langle N_{\text{coll}} \rangle} \frac{\mathrm{d}^2 N_{AA}/\mathrm{d}p_T}{\mathrm{d}^2 \sigma_{pp}/\mathrm{d}p_T} = \frac{1}{T_{AA}} \frac{\mathrm{d}^2 N_{AA}/\mathrm{d}p_T}{\mathrm{d}^2 \sigma_{pp}/\mathrm{d}p_T} \tag{1.23}$$

The T_{AA} is known as the nuclear overlap function and can be calculated from the Glauber model as discussed in Section 1.3.3. The description about experimentally determining the N_{coll} and it's relation to the collision centrality is described in Section 4.2.

The R_{AA} serves as a measure of medium effects, such that for $R_{AA} = 1$ heavy ion collisions can be treated as a superposition of multiple pp collisions. Colorless probes which are not expected to strongly interact with the medium are expected to have a $R_{AA} \approx 1$. Initial state quarks and gluons producing in the hard scattering on the other hand, interact with the QGP and are detected as a jets with a lower energy as discussed in the earlier sections. The R_{AA} for a colorless probe (photon) and for inclusive jet measurement is shown in Fig. 1.24 [75, 76]. The contrast showing a $R_{AA} \approx 1$ for photon

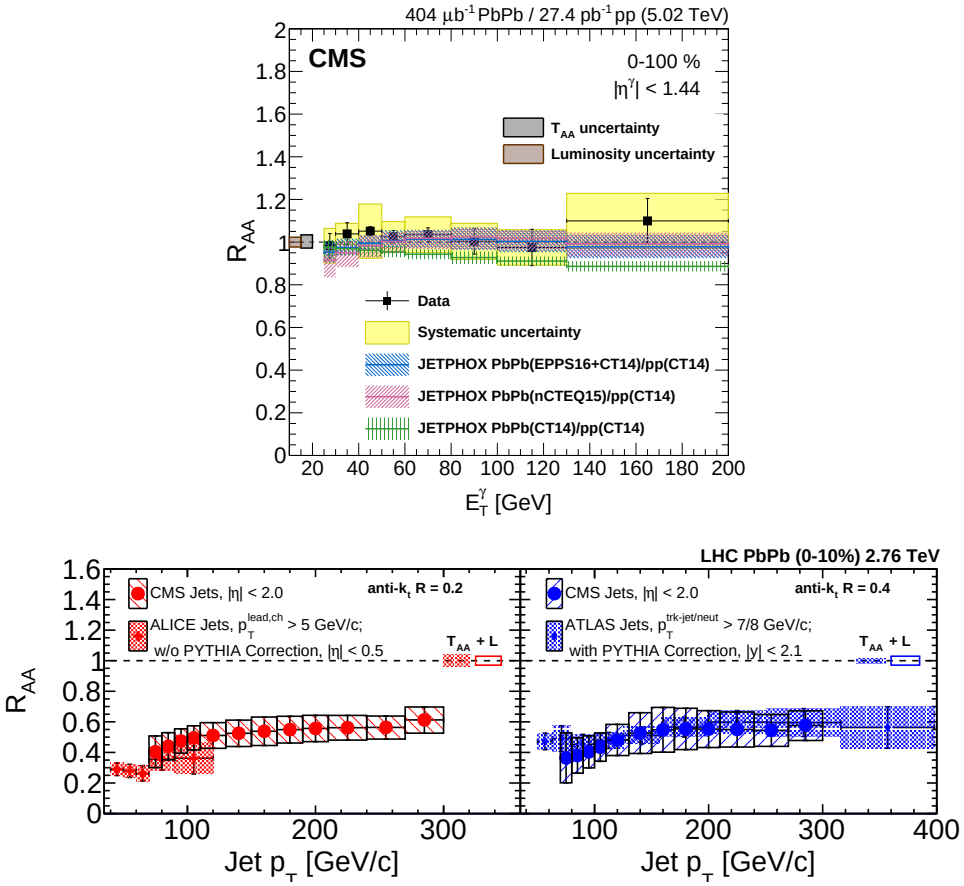


Figure 1.24: R_{AA} for photons (upper) [75] and inclusive jets (lower) [76] showing no modification for photons and a suppression in jets

and a suppression in the jet R_{AA} is seen as expected.

Another observable that allows for a direct measure of the degree of jet quenching is the momentum imbalance (x_J). It is described by ratio of momenta of back to back jets as shown in Fig.1.25 [77]. In inclusive jet measurements, the depletion of the balanced jets in central PbPb collisions is understood by means of the path length, such that one of the jets traverses more of the QGP medium than the other. This leads to greater energy loss and more imbalanced jets. The magnitude of energy loss is a very broad spectrum and the mechanism is not fully understood, for instance the peak seen in the x_J distribution. The momentum imbalance is an observable highly sensitive to effects

of fluctuations in the initial geometry and the energy loss process.

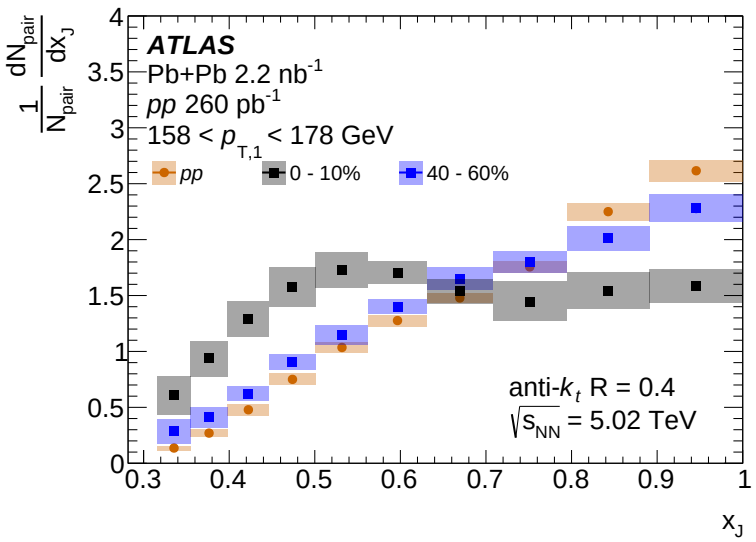


Figure 1.25: Momentum imbalance in pp collisions compared to central and peripheral PbPb collisions [77]. A shift in the distribution is seen for central PbPb collisions compared to peripheral or pp collisions as a signature of jet quenching.

Jet substructure

Earlier studies of jet quenching considered jets as monolithic objects. As we have discussed in the earlier sections, the radiation pattern of the hard parton is affected by the medium, undergoing multiple scatterings depending on the energy scale. Jet substructure offers various tools to probe the history of the jet clustering as discussed in Section 1.2.5. Jet measurements in heavy ion collisions capture the full dynamics of jet quenching through the construction of various observables. However, determining the properties of the QGP from the modifications of jets is a non trivial extrapolation. The understanding of the QGP gained from modifications to jets is limited by the precise mechanisms of the interactions of the jet with the medium and the predictive power of the different theoretical frameworks. Also, the experimental measurement of medium modified jets presents a challenge due to the fact that any measurement is performed after the jet has propagated through the entire QGP and isolating the initial state modi-

fications to those effects following the QGP is not straightforward. The large fluctuating background from the UE in heavy ion collisions also poses a challenge in studying the hard scattering. Jet substructure measurements have been introduced as a method to combat some of the aforementioned issues. It has various applications such as tuning MC generators to describe the data considering the evolution of the jet or tagging boosted particles such as heavy quarks, to testing the standard model [22].

The jet shower in heavy ion collisions is expected to factorize into an early vacuum shower, created before the medium formation, followed by a medium-modified showering process [54]. It is useful to describe the cascade of emissions in a parton shower in stages, (i) the early parton shower and (ii) medium interactions. The early parton shower consist of emissions of the parton with small formation time (t_f) such that the emissions happen before interactions with the medium ($t_f \ll L$, where L is the length of the medium). Thus, the early parton shower does not see the medium and behaves similar to the parton showering in vacuum. Emissions with larger formation time ($t_f \geq L$) interact with the QGP medium undergoing multiple scatterings and changing their radiation pattern. This is the stage where the emissions are resolved into color charges in their interaction with the QGP, depending on the medium resolution length. These can be summarized as formation time arguments where the early parton shower defines the structure of the jet (broad or narrow) and hence, the degree of quenching that the jets undergo. Jet substructure observables provide insight into the properties of the jet showering.

As discussed in Section 1.2.5, jet substructure observables are classified based on the clustering history or energy flow and inclusive jet measurements of these observables have provided the most consistent results in the understanding of the QGP. Experimentally, light quarks and gluons are the dominant source of inclusive jets. Due to the high rate of inclusive jet events in hadronic collisions, they can be studied extensively

over a wide kinematic range. A summary of measurements across both categories of jet substructure observables is provided henceforth.

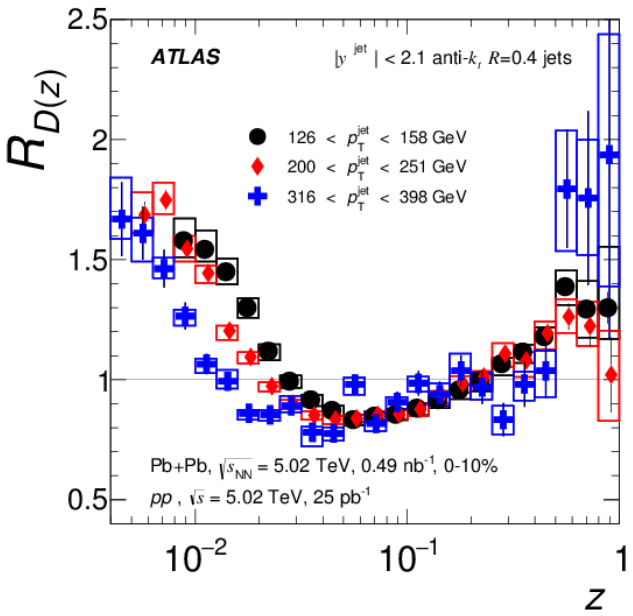


Figure 1.26: Ratio of jet fragmentation function showing differences for low momentum fraction z and consistency at high z for different jet p_{T} ranges [78]

The characteristic evolution of the parton shower determines the degree of energy loss, primarily dependent on the formation time arguments and color coherence. Gluon jets are expected to be broader compared to quark jets and this difference is characterized by the fragmentation function of the jet. In Fig.1.26, z is the distribution of the momentum fraction of a charged particle in the jet [78]. At high-z the ratio of fragmentation function in pp and PbPb collisions are consistent for different p_{T} ranges and there is an excess of high-z particles. An explanation of this is a selection bias as high-z particles are caused due to harder fragmentation and are thus, quenched less. It could also be understood as gluon jets being quenched more than quark jets due to hard fragmentation on average in quark jets.

Jet shape observables provide insight into the angular distribution of energy in the jets. Figure. 1.27 shows two measurements of angularities described in eq. 1.17 [79] namely,

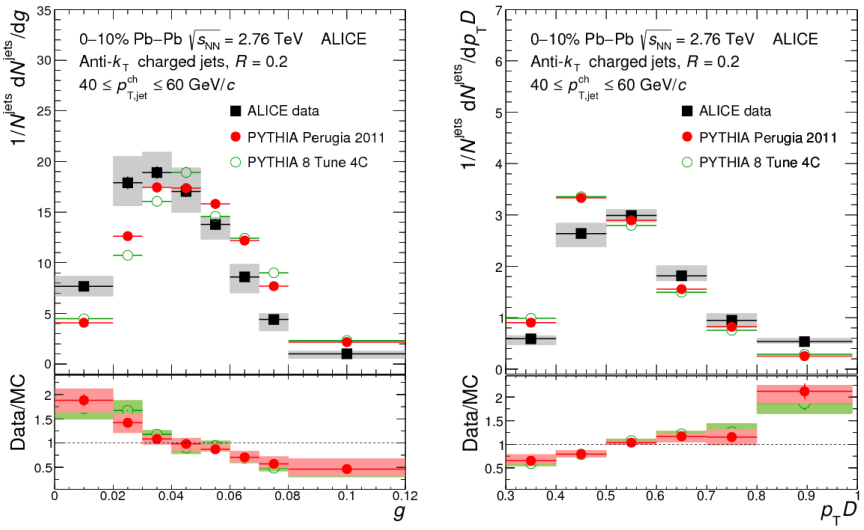


Figure 1.27: Measurements of the angularity observables, the jet girth g(left) and momentum dispersion $p_T D$(right) in PbPb collisions compared to simulation [79]. There is a characteristic suppression of broader jets and an enhancement of narrower jets

the jet girth (g) and momentum dispersion ($p_T D$). The distributions for PbPb collisions for these observables show that measured quenched jets are narrower and have harder fragmentation than in simulation. This could also be interpreted as a sign of selection bias where broader jets are quenched more due to their softer fragmentation.

Another tool in jet substructure measurements, is Soft Drop grooming. By studying the clustering history a procedure known as grooming is performed which can remove large angle and soft radiation yielding a sufficiently hard splitting in the jet, making the jet less prone to nonperturbative effects such as pileup or the UE. Through this process of grooming various substructure observables can be defined either using the positions and momenta of the jet constituents directly or through the use of the clustering history.

Experimental results from heavy ion collisions provide evidence for medium-induced jet modifications [78–91]; recent reviews can be found in Refs. [92, 93]. Figure. 1.28 shows the measurements of the groomed momentum balance (z_g) and groomed jet ra-

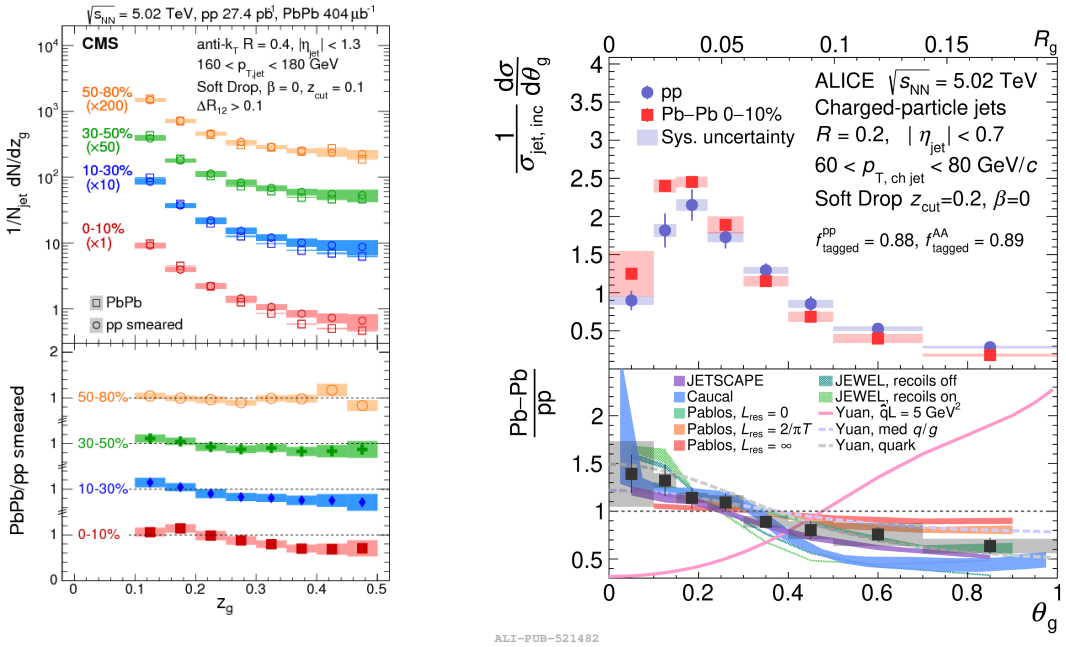


Figure 1.28: Measurement of groomed momentum balance z_g by CMS (left) [80] and the groomed jet radius R_g by ALICE (right) [81] showing a suppression of large angle splittings

dius (R_g). The z_g in vacuum is connected to the splitting function displaying a universal behavior of $1/z$, while in PbPb collisions the evolution of the jet shower is intertwined with the medium induced radiation. The measurement shows a suppression of large angle splittings while enhancing small angle splittings.

The other observable that has been used to study the modifications of the radiation pattern of the jet is R_g, defined as the rapidity–azimuth (y–ϕ) distance between the two subjets obtained using the soft-drop grooming algorithm [32, 94].

Measurements of R_g at the RHIC and LHC show a narrower jet substructure in heavy ion collisions compared with pp collisions for jets reconstructed with the same jet transverse momentum p_T [81, 90, 91, 96]. This narrowing of the hard intrajet angular distributions measured via R_g could be a manifestation of the medium resolution length of the QGP. However, before investigating such a connection, one should assess the po-

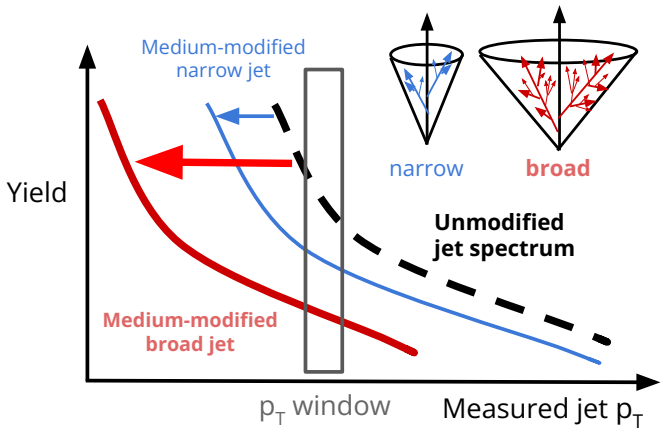


Figure 1.29: Schematic diagram of the potential selection bias due to jet energy loss that may occur when selecting jets based on the their p_T. Broader structures are expected to be more quenched (thicker red line and arrow), whereas narrower structures are expected to be quenched less (thinner blue line and arrow). Combined with the steeply falling jet p_T spectrum, this can lead to a preferential selection of narrow jets in a given jet p_T interval, as indicated by the vertical rectangular box. The dashed curve represents the jet p_T spectrum in the absence of medium-induced jet modifications. [95]

tential selection bias discussed in the earlier observables that emerges when comparing PbPb and pp jets with the same reconstructed p_T [97–99]. Since jet energy loss may fluctuate jet-by-jet and the jet p_T spectrum is steeply falling, a given jet p_T interval is preferentially populated by jets that are less quenched. Selecting jets with large R_g values might preferentially isolate early vacuum radiation patterns that are broad in angle and thus interact more frequently with the QGP, which leads to stronger jet energy loss and to migrations to lower jet momenta, as illustrated schematically in Fig. 1.29 [95]. Thus, to interpret the modifications of the radiation pattern of the jet in terms of the emergence of other physical scales, such as the medium resolution length, selection bias effects must be understood first.

1.4 Constraining jet quenching with photon tagged jet substructure

Prompt photons produced in quark-gluon Compton scattering and quark-antiquark annihilation processes, in which their momentum is balanced by a recoiling jet, provide an ideal topology to study medium-induced jet modifications. A high-momentum photon does not interact strongly with the QGP [100–104], so its p_{T} can be used as a proxy for the p_{T} of the recoiling parton that initiates the jet shower. By selecting on the photon p_{T}, selection biases associated with jet energy loss should be reduced. Prompt photons can be studied using photon isolation techniques [105, 106]. Isolated photon+jet events have been used to quantify the jet energy loss via the measurement of the momentum imbalance, defined as the ratio of the p_{T} of the highest-p_{T} jet and the photon momentum $x_{\gamma \mathrm{j}} \equiv p_{\mathrm{T}}^{\mathrm{jet}} / p_{\mathrm{T}}^{\gamma}$ [75, 84, 88]. The results show a stronger momentum imbalance in PbPb collisions than in pp collisions due to the jet energy loss that occurs in the former. Fragmentation functions have been measured in isolated photon+jet events in PbPb collisions, showing a suppression of hard jet constituents and an

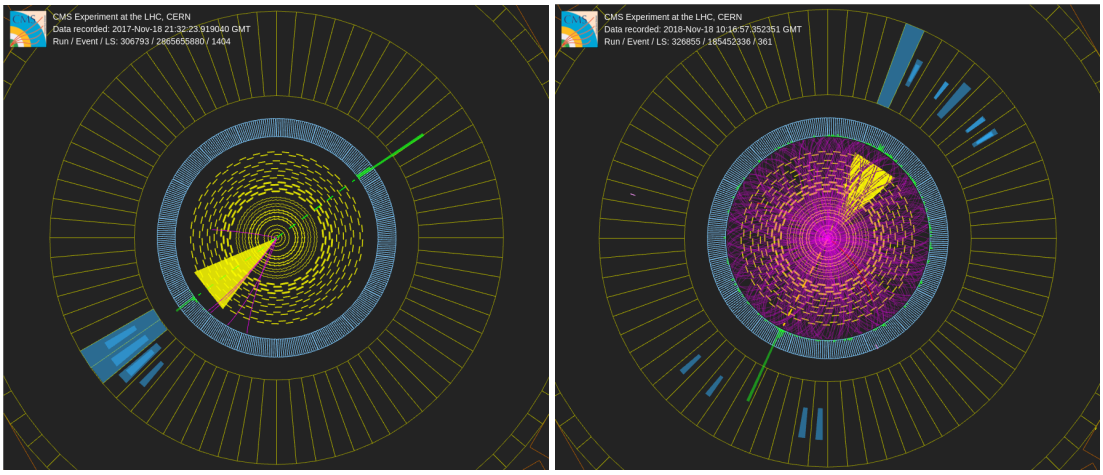


Figure 1.30: Event display for a photon tagged jet event in pp (left) and PbPb (right) collisions. The green bar represents the photon with the energy deposit primarily in the ECAL. The jet is represented by the yellow cone consisting of tracks (in purple), ECAL energy deposits (in green) and HCAL energy deposits (in blue)

enhancement of soft constituents relative to pp collisions [107].

Figure. 1.30 shows a characteristic photon tagged jet event in pp and PbPb collisions. By considering the $x_{\gamma \mathrm{j}}$ in both these systems, a fair comparison is possible due to the photon momentum not being affected strongly due to the QGP. The PbPb environment is busier than the pp as can been seen due to the underlying event creating multiple tracks and energy hits in the calorimeters. The reconstruction of physics objects considering these different environments is discussed in greater detail in Chapter 3.

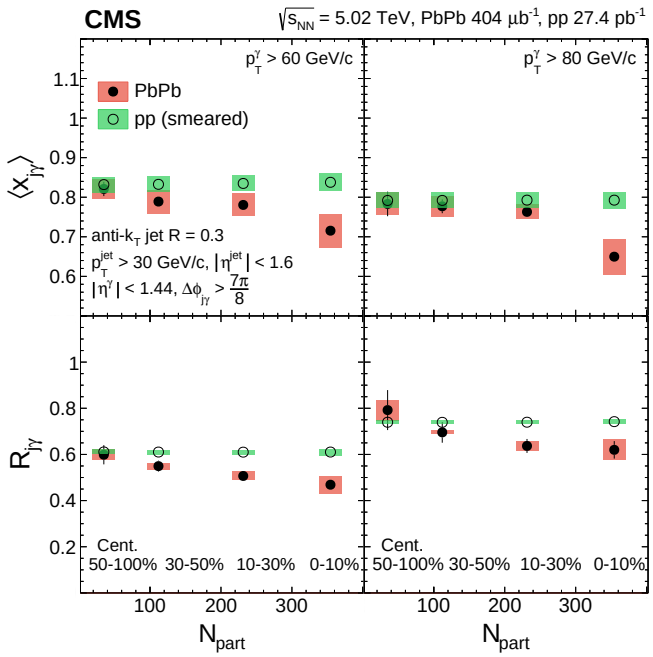


Figure 1.31: The centrality dependence of $\langle x_{\gamma \mathrm{j}} \rangle$ (upper) and $R_{j\gamma}$(lower) of photon+jet pairs normalized by the number of photons for PbPb (full markers) and smeared pp (open markers) data[84]. A suppression is observed for central PbPb collisions for both observables

A measurement of the jet quenching performed in photon+jet events is shown in Fig.1.31[84]. Average momentum imbalance $\langle x_{\gamma \mathrm{j}} \rangle$ and the average number of associated jets per photon $R_{j\gamma}$ as a function of centrality both show a suppression of central (0–10%) PbPb yield compared to smeared pp. This is consistent with the expectation that the limited size of the jet cone in PbPb collisions leads to greater energy loss due to greater average

path length that the parton needs to traverse in central PbPb collisions. A measurement comparing the nuclear modification factor for inclusive and photon tagged jets is shown in Fig.1.32. The R_{AA} of jets is expected to be lesser than 1 as seen in Fig. 1.24 but a relative suppression is seen for inclusive jets with respect to photon tagged jets as a function of the jet momentum. An inclusive measurement as a function of R_g is consistent with previous studies of narrowing of jet substructure. But the selection bias described earlier and the difference seen in inclusive and photon tagged jets is expected to affect the substructure as well.

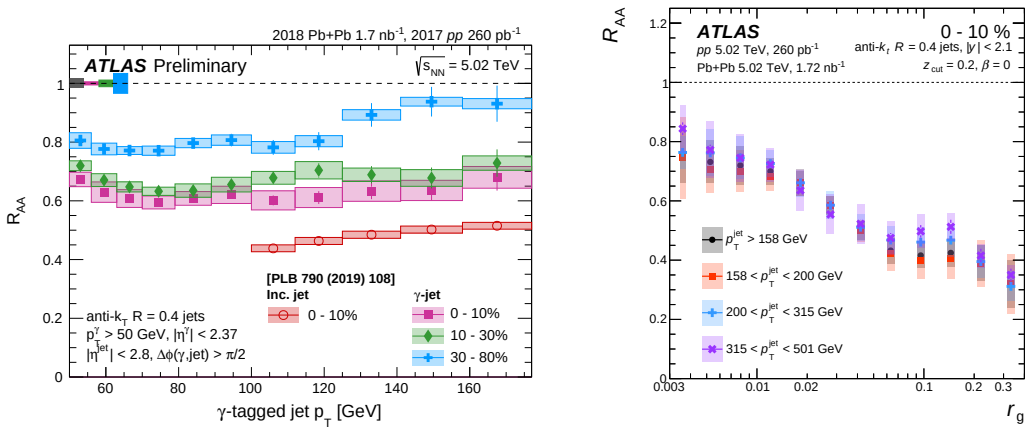


Figure 1.32: Inclusive vs photon tagged jet R_{AA} as a function of photon tagged jet p_T(left) [108, 109] and inclusive jet R_{AA} as a function of R_g (right) [96] measured by ATLAS. The difference in the R_{AA} between inclusive and photon tagged jets shows a clear dependence on the color factor for parton-QGP interactions. This is expected to modify the jet substructure R_g for photon tagged jets

The goal of this thesis is to study the jet substructure in photon-tagged jet events in an effort to mitigate the selection bias due to color factors present in inclusive jet measurements. The quenching is not well controlled in inclusive jets which is a strength of colorless probes such as photons. By using the photon p_T as a proxy for the parton momentum, the jet substructure in photon tagged jets can be studied for very quenched jets. This allows us to probe the fundamental properties of the QGP such as the medium resolution length, and broadly allowing a better understanding of QCD in the high tem-

perature, low net baryon density regime. Jet observables such as ungroomed jet girth g and groomed jet radius R_g are studied in the CMS detector providing complementary information about the substructure comparing pp and PbPb collision data obtained in 2017 and 2018 respectively as part of the Run 2 data collection.

This thesis is structured to provide all the details necessary in the context of these studies. In Chapter 1, a discussion about the current measurements and the physics motivation for this analysis has been discussed. Chapter 2 provides details about the CMS detector used in the collection of data. Detailed information about reconstructing physics objects such as photons and jets used in this analysis from the detecter information is provided in Chapter 3. Chapter 4 and 5 discuss the analysis procedure followed and the checks performed at each stage to ensure robust scientific results, respectively. The results of the analysis performed in this thesis are presented in Chapter 6 [95].

Chapter 2

Experimental setup

The experimental studies of particle physics have made significant progress over the previous century, but fundamentally the study involves colliding various species and observing the results. The long history of particle accelerators at CERN starts from the *Synchro-cyclotron* with the discovery of the electron decay of the pion [110], through the *Proton Synchroton (PS)*, *Gargamelle bubble chamber*, to colliders such as *Intersecting Storage Rings (ISR)*, *Super Proton Synchrotron (SPS)* to arrive at the present, the *Large Hadron Collider (LHC)* [111]. The work of this thesis focuses on the physics programs at LHC, specifically using the CMS detector which is the focus of this chapter.

2.1 LHC

The LHC is currently the largest particle collider in the world with a circumference of 27 km, located between the French and Swiss border at a depth of 50 to 175 m underground. The LHC consists of two intersecting rings, where hadron beams circulate in clockwise (anti-clockwise) direction moving close to the speed of light. Hadrons (protons and ions) are accelerated to these high energies through an accelerator complex, maintained by superconducting electromagnets that operate at -271.3°C. The protons and ions are accelerated through progressively increasing energies in an injection chain as shown in the CERN accelerator complex in Fig. 2.1.

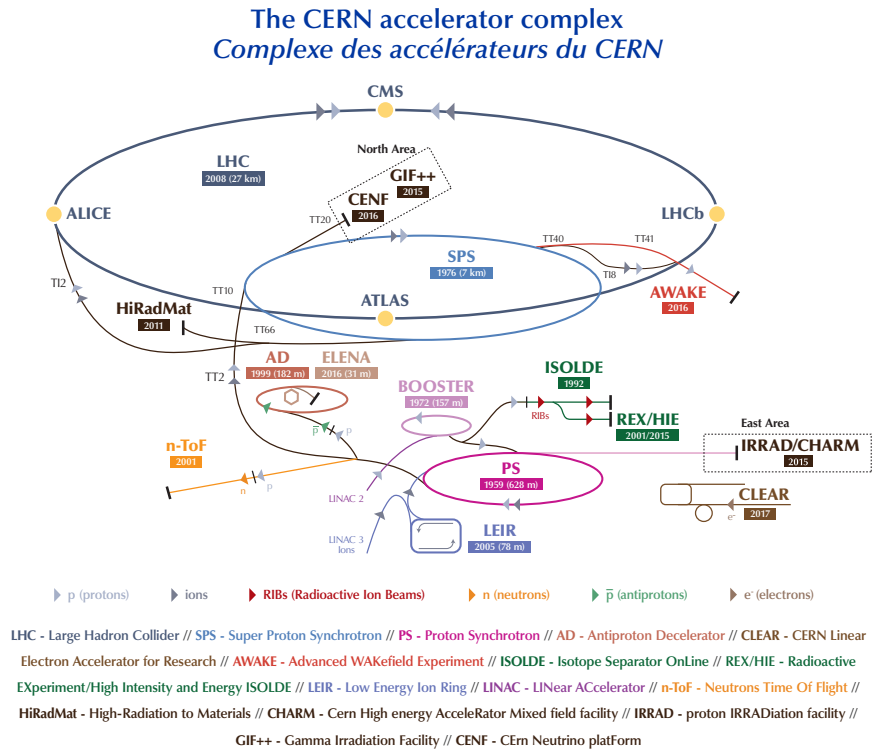


Figure 2.1: Schematic diagram of the CERN accelerator complex [112]

Protons are first extracted from hydrogen gas with an electric field and pass through the linear accelerator, LINAC2 to be accelerated to 50 MeV using radio frequency (RF) cavities. The protons are then injected into the Proton Synchrotron Booster (PSB) to be accelerated up to 1.4 GeV followed by the Proton Synchroton (PS). At the PSB the protons are grouped into bunches separated in time by 25 ns by the time they are injected into the Super Proton Synchrotron (SPS).

In the case of ions, (^{208}Pb which is the focus of this thesis), due to the high charge the electrons are stripped away at each stage. In the case of ^{208}Pb, 29 weakly bound electrons are stripped away at the Electron Cyclotron Resonance ion source to inject $^{208}Pb^{29+}$ to the LINAC3 with an energy of 2.5 keV per nucleon. The LINAC3 further accelerates the ions to 4.2 MeV per nucleon stripping away more electrons to inject $^{208}Pb^{53+}$ to the Low Energy Ion Ring (LIER). The ions are then grouped into bunches

and accelerated up to 72 MeV and stored in the PS. The ions are then fully stripped of electrons before injection the SPS.

The SPS takes both proton and/or ion beams, accelerates the batches and injects them into the 2 LHC beam pipes in opposite directions. Protons are brought up to 450 GeV or Pb nuclei to 177 GeV per nucleon at the SPS. At the LHC, they are accelerated to their final energy of 5.02 TeV per nucleon pair (2.51 TeV per nucleon for each beam) for the analysis performed using the data collected in 2017 for pp collisions and 2018 for PbPb collisions studied in this thesis. Other runs inject various species such as Xe and up to 13 TeV for protons, with the machine designed for a center of mass proton collision energy of 14 TeV in future runs.

The beams cross at four interaction points that host the four largest CERN experiments namely,

- *A Large Ion Collider Experiment* (ALICE) [113]: The ALICE experiment primarily focuses on the study of nuclear matter at high densities and temperatures. It was conceived as the primary experiment in the study of the QGP in heavy ion collisions such as ^{208}Pb and has a strong advantage in its ability to reconstruct and identify low-p_{T} particles such as pions and kaons. The detector with its 18 sub detectors mainly reconstructs the bulk of charged particles and electromagnetic probes at mid rapidity, with a focus on muons in the forward region. This provides complementary information to lower energy QCD studies from the SPS and Relativistic Heavy Ion Collider (RHIC).
- *A Toroidal LHC ApparatuS* (ATLAS) [114]: ATLAS is the largest detector at the LHC and one of two multi-purpose detectors. A primary goal of ATLAS was to detect the Higgs boson and other heavy particle decays in multiple channels. The versatility of the detector allows for a wise range of studies including participation in the heavy ion program of the LHC to comple-

ment those of the ALICE experiment. It also has a similar physics reach as the CMS experiment providing redundancy through different experimental methods and was fundamental to have confidence in the Higgs boson discovery. The detector is shaped like a cylinder containing a magnet in the shape of a toroid, an inner tracker, and electromagnetic and hadronic calorimeters and a muon spectrometer.

- *Compact Muon Solenoid* (CMS) [115]: The CMS detector is the other multi purpose detector, complementing ATLAS and is also part of the heavy ion program. This thesis focuses on a measurement using the CMS detector and Section 2.2 is dedicated to discussing its components in detail.
- *LHC beauty* (LHCb) [116]: The aim of the LHCb experiment is on investigating CP symmetry and baryogenesis i.e. the slight differences between matter and antimatter through the study of b quarks. It is a forward single-arm spectrometer dedicated to heavy flavor physics. It has a particular advantage in momentum and vertex resolution over the other experiments to differentiate protons, kaons and pions at the cost of a limited acceptance in rapidity. In addition to pp and heavy ion collisions, the LHCb can operate in fixed target mode by injecting nuclear gas (He, Ne, Ar) into the beam pipe at the interaction point.

Other experiments as part of the LHC include the *LHC forward* (LHCf) [117], *TOTal cross-section, Elastic scattering and diffraction dissociation Measurement at the LHC* (TOTEM) [118] and *MOnopole and Exotics Detector at the LHC* (MOEDAL) [119].

2.2 CMS

The CMS detector is one of the two general purpose detectors at the CERN LHC. At 21 m long, 15 m in diameter and 14 000 tonnes, it has similar capabilities as ATLAS in a much more compact design. The subdetectors of the CMS are arranged in a cylindrical fashion as shown in Fig. 2.2 around the beam axis giving it a classic "onion-like" structure making it quasi-hermetic (covering almost a 4π solid angle) [120, 121]. It is designed to trigger on and identify electrons, muons, photons along with charged and neutral hadrons [122–124]. The analysis presented in this thesis uses proton-proton (pp) and lead-lead (PbPb) collision data collected by the CMS experiment in 2017 and 2018, respectively. The pp and PbPb data samples, both at a nucleon-nucleon center-of-mass energy of 5.02 TeV, correspond to integrated luminosities of 301 pb^{-1} and 1.7 nb^{-1}, respectively, which are known with a precision of 1.5% and 1.9% [125–127].

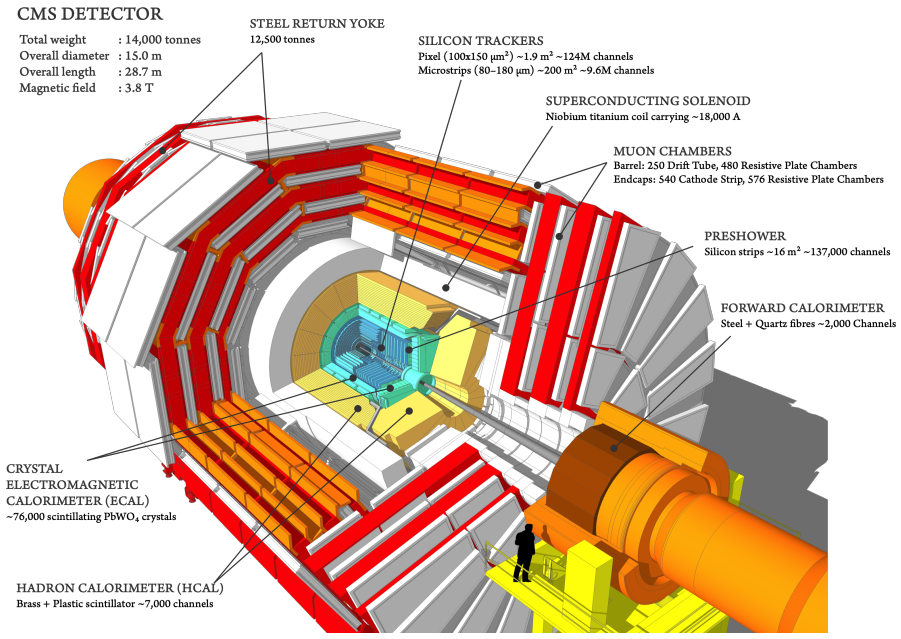


Figure 2.2: Schematic drawing of the CMS detector [128]

2.2.1 Coordinate system

The coordinate system used in accelerators have many standard variables and notations in common. A general summary of the coordinate system with a focus on the CMS experiment is provided here as a guideline to the reader. The cartesian coordinates x, y, and z in the detector are defined such that the z- axis is along the beam direction, and the y-axis points to the sky with the x-axis pointing inward in the case of circular colliders such as the LHC. A spherical coordinate system is commonly used to describe the position of particles or detector components with respective to the collision vertex as shown in Fig. 2.3. The polar angle θ is the angle between a given particle's three-momentum (p) relative to the z-axis, while the azimuthal angle ϕ is measured from the x-axis on the transverse plane. The transverse momentum $p_{\mathrm{T}} = \sqrt{p_x^2 + p_y^2}$, where p_x and p_y are the components of the three-momentum of the particle along the x-axis and y-axis. The p_{T} is a Lorentz boost invariant quantity, used frequently as a proxy for the scale of the scattering process.

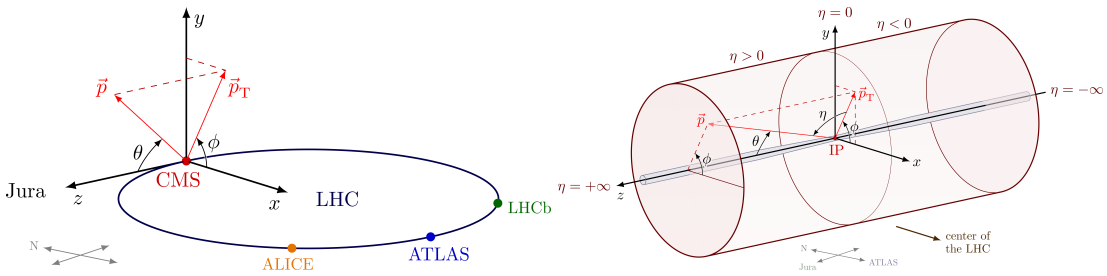


Figure 2.3: The conventional 3D coordinate system at the CMS detector [129]

Another quantity known as the rapidity y (distinguished from the cartesian coordinate) provides a link between the angle θ and the particle's momentum defined as,

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right) = \tanh^{-1} \frac{p_z}{E} \tag{2.1}$$

where E is the total energy and p_z is the longitudinal momentum. The rapidity y is an interesting quantity as it is additive under logitudinal Lorentz transformations, imply-

ing that the difference of rapidities of two particles is invariant under a longitudinal boost. The rapidity ranges from 0 for particles in the transverse plane, referred to as mid-rapidity to $\pm\infty$ along the beam directions. In experiments, it is however more convenient to only use the beam-momentum angle instead through the pseudorapidity η, as measuring the particle's energy and momentum along the beam direction can be hard to measure precisely. The pseudorapidity is then defined as,

$$\eta = \frac{1}{2}\ln\left(\frac{|p| + p_z}{|p| - p_z}\right) = -\ln\left(\tan\frac{\theta}{2}\right) \tag{2.2}$$

In the case relativistic particles $y \approx \eta$ and they are equal for massless particles, thus η is also Lorentz invariant under longitudinal boosts. Figure 2.4 shows the link between the angle θ and pseudorapidity η following this relation.

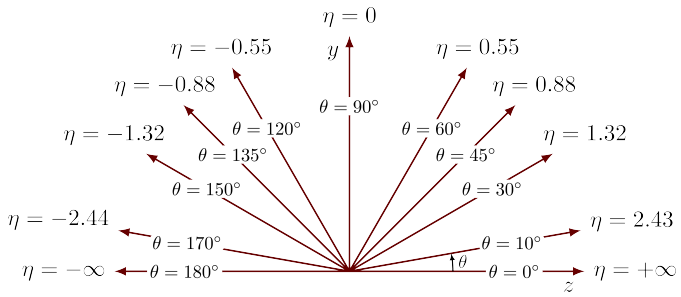


Figure 2.4: Pseudorapidity (η) on a 2D coordinate axis showing the relationship with angle [129]

Another commonly used quantity is the distance measure between two particle in the (η, ϕ) plane. For 2 particles with angular coordinates (η_1, ϕ_1) and (η_2, ϕ_2) it is defined as,

$$\Delta R = \sqrt{(\eta_1 - \eta_2)^2 + (\phi_1 - \phi_2)^2} \tag{2.3}$$

where $(\phi_1 - \phi_2)$ is taken in the domain $[-\pi, \pi]$.

2.2.2 Magnet

The CMS detector owes part of its name to the quasi-uniform magnetic field of 3.8 T which is about a 100 000 times the strength of the Earth's magnetic field, provided by the 6 m diameter superconducting solenoid. The presence of a strong magnetic field is a crucial component in the reconstruction and identification of particles. With the magnetic field along the beam direction, charged particle trajectories are bent while traversing the detector. The CMS magnet system consists of the superconducting coil, the vacuum tank, and the magnet return yoke shown in a longitudinal cross section in Fig. 2.5. The figure shows a prediction of the magnetic field distribution and the field lines.

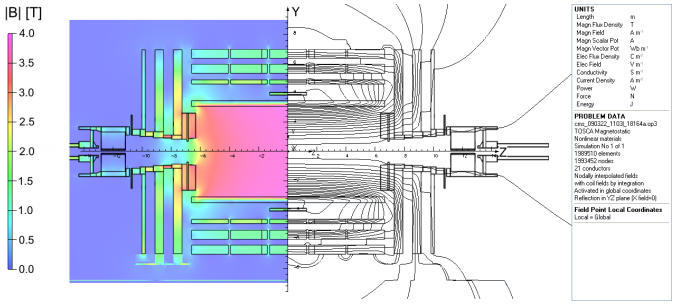


Figure 2.5: Prediction of $|B|$(left) and field lines (right) on a logitudinal section of the CMS detector at a central magnetic flux density of 3.8 T [130]

The solenoid consists of four layers of Rutherford type, Niobium-Titanium superconducting strands, which were extruded with pure aluminium and mechanically reinforced with an aluminium alloy. The superconducting coil is thermally insulated in vacuum and operates by letting electric currents flow without resistance. The "return yoke", thus named due to its responsibility of returning the magnetic flux is a 12 sided iron structure surrounded by that surrounds the magnetic coils.

The magnetic field is such that charged particle trajectories are bent in the transverse x-y plane, while z-component of the momentum remains relatively unaffected. In radial coordinates, this is seen as a smearing in ϕ, with relatively no effect in the η direction.

The bending direction and magnitude are significant factors in determining the charge and the particle momenta.

2.2.3 Tracker

The detector part closest to the beam is the tracker system. It is primarily divided into the Tracker Inner Barrel (TIB), Tracker Outer Barrel (TOB), Tracker Inner Disk (TID) and Tracker End Cap (TEC) as shown in Fig. 2.6 for an overall coverage of $|\eta| < 2.5$. The tracking system is designed to have high granularity to record and reconstruct particle trajectories through registering electromagnetic "hits" in the detector modules. When charged particles pass through the tracker, they interact with the silicon atoms creating electron-hole pairs. The free electrons then drift in an electric current which is amplified and detected. A key aspect of the tracker is to be as lightweight and thin as possible to reduce interaction with the particle while detecting it hits. The tracker modules are categorized into pixel detectors and silicon strip detectors. The pixel detectors make up the 4 closest layers in the barrel, staring from 2.9 cm from the collisions and 3 forward pixel disks for a total of 1440 modules. Each hit in the pixel detectors is reconstructed with a granularity of 10 (20-40) microns in the transverse (longitudinal) directions. The silicon strip detector layers surrounds the pixel layers consisting a total of 15 148 modules extending out to a radius of 110 cm.

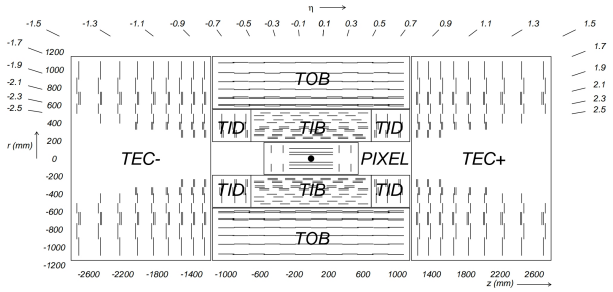


Figure 2.6: Schematic cross section through the CMS tracker. Each line represents a detector module. Double lines indicate back-to-back modules which deliver stereo hits [115]

2.2.4 Electromagnetic calorimeter (ECAL)

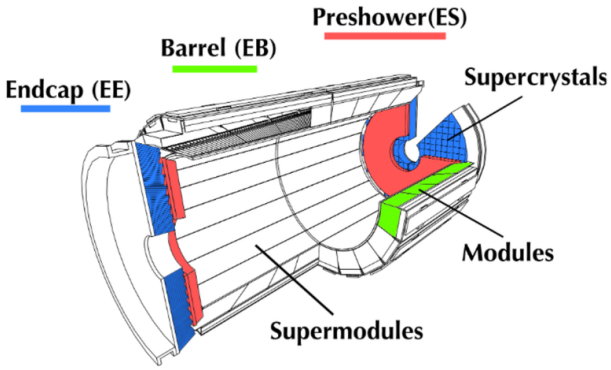


Figure 2.7: Schematic ECAL layout highlighting the Barrel (EB), Endcap (EE) and Preshower (ES) detectors [131]

The Electromagnetic Calorimeter (ECAL) is one of the calorimeter systems in the CMS detector. Calorimeters unlike the tracker system are designed to measure the energy of the particles by fully absorbing them such that they do not propagate to further outer systems in the CMS detector. The ECAL is designed with the goal of detecting all electrons and photons. The ECAL is made up of lead-tungstate crystals, chosen for its short radiation length and high density, resulting in a finely granular and compact calorimeter as shown in Fig. 2.7. The ECAL is composed of the barrel (EB), Endcap (EE) and Preshower disks (ES) to cover a total pseudorapidity range of $|\eta| < 3$. The transverse size of the crystals is $2.2 \times 2.2\,\mathrm{cm}^2$ in EB and $2.86 \times 2.86\,\mathrm{cm}^2$ in EE. They are organized into 36 supermodules in the barrel each containing 1700 crystals, with the endcap consisting of 3662 crystals on each side with the preshower detector in front of it with 4288 sensors. High energy charged particles passing through the ECAL deposit energy in the crystals through various processes, including electron-positron pair production and bremsstrahlung for photons and electrons as discussed in further detail in Section 3.2.2. This results in the crystals emitting scintillation light which is detected by the photodetector. In the barrel, Avalanche PhotoDiodes (APD) are mounted in pairs at the back of each crystal while in the endcap Vacuum PhotoTriodes (VPT) are used. The

endcaps also contain the preshower detectors which consist of two layers of lead followed by silicon strip detectors to identify neutral pion decays which enhances photon identification [131]. The crystals are kept at 18°C using a water flow cooling system, since the response of the crystals and the signal amplification depend on the temperature. The transparency of the crystals to scintillation light is also affected by radiation and laser pulses are used to monitor the radiation induced transparency variations.

2.2.5 Hadronic calorimeter (HCAL)

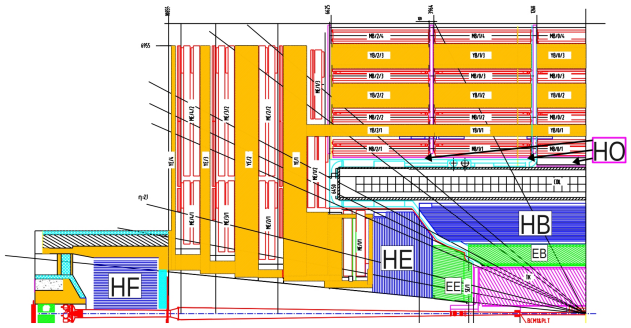


Figure 2.8: Longitudinal view of one quadrant the CMS detector showing the locations of the hadron barrel (HB), endcap (HE), outer (HO) and forward (HF) calorimeters [132]

The Hadronic Calorimeter (HCAL) measures the energy, position, and time of arrival of hadronic particles. This is achieved through the use of successive layers of absorber and scintillating materials. The HCAL and ECAL form a complete calorimetery system for the measurement of jets as discussed in Section 3.3 along with missing transverse energy. The HCAL consists of 4 sections as shown in Fig. 2.8 [132]. The barrel (HB) and endcap (HE) components are located within the magnetic field of the CMS detector. On the other hand, the outer barrel (HO) and forward (HF) subsystems are located outside the solenoid within the return flux, where the magnetic field is significantly smaller. The HB and HE utilize brass absorber and plastic scintillator, with the resulting light measured by Hybrid PhotoDiode (HPD). The HO complements the HB and uses Bicorn BC408 scintillator plates and Silicon PhotoMultipliers (SiPM) to detect photons.

The HF located at very forward rapidities covering $2.9 < |\eta| < 5.2$ experiences unprecedented particle fluxes and extremely high charged hadron rates. The active material is made of quartz to resist the high radiation rates near the beam along with a steel absorber. The fibres collect the Cherenkov radiation emitted by charged particles, thus primarily detecting electromagnetic showers. The transverse energy deposited in the HF is used to measure the luminosity and especially in heavy ion collisions the global event activity defined by the centrality is determined using the HF as described in Section 4.2.

2.2.6 Muon chambers

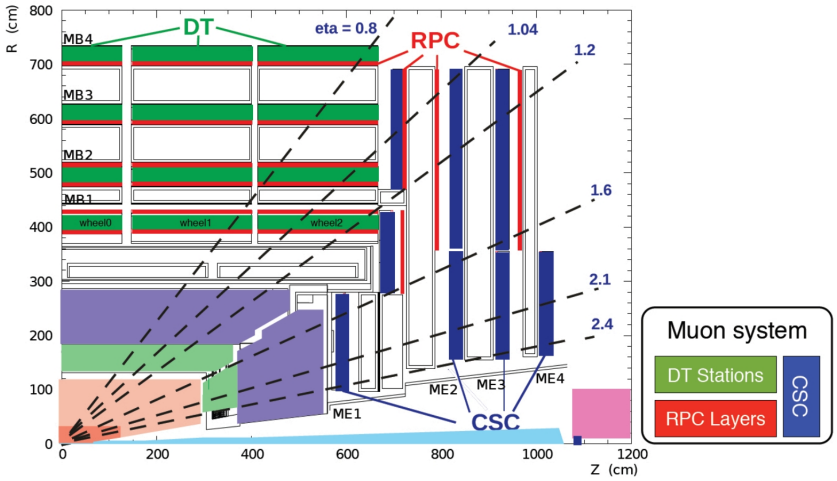


Figure 2.9: Longitudinal layout of one quadrant of the CMS detector showing the locations of the drift tubes DT (green), cathode strip chambers CSC (blue) and resistive plate chambers RPC (red) [133]

The muon chambers consisting of Drift Tubes (DT), Cathode Strip Shambers (CSC) and Resistive Plate Chambers (RPC) are part of the outermost subdetectors of CMS as shown in Fig. 2.9. The muon detection in CMS is one of the defining features of the detector. As muons penetrate the dense calorimeteric materials which absorb all the other particles, the muon chambers solely focus on the detection of muons. The DTs measure the position of muons accurately in the barrel region, while the CSCs are used

to measure the position of the muons in the endcap disks where the magnetic field is not uniform. The CSCs and RPCs are fast detectors suitable to provide a muon trigger system.

2.2.7 Trigger

The CMS detector employs a trigger system to select collision events of interest to the physics community. The beams at the LHC arer designed to have a crossing frequency of 40 MHz or once every 25 ns in the case of pp collisions. Each of these beam crossings consist of multiple interaction events. Most of these events are due to long-distance physics interactions collectively called "minimum-bias" (MB). These MB events involve low momentum transfers in their interactions and have a large cross section making them easier to produce and study. While MB events are interesting, the primary interest in high energy collisions is studying rarer process such as the production of particles at high p_{T} and masses. The machine is also limited in the events that can be stored for later offline analysis motivating the need for a trigger system to select interesting events for offline storage. A two-level trigger system is employed to efficiently select

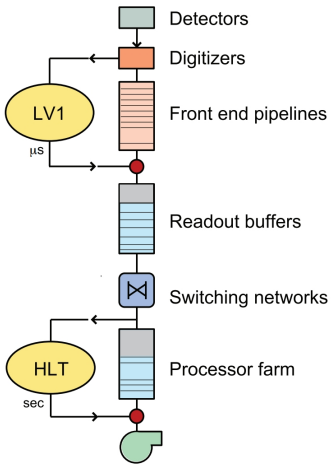


Figure 2.10: Schematic overview of the CMS trigger system showing the L1 and HLT processes [134]

events for offline analysis as shown in Fig. 2.10. These two selection stages are called Level 1 trigger (L1) and the High-Level Trigger (HLT), where we select the events given a maximal rate of 100 kHz for L1 and 100 Hz for HLT.

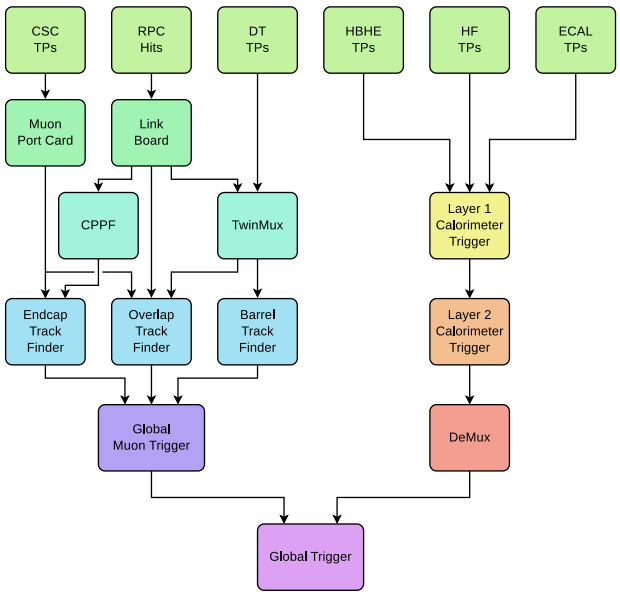


Figure 2.11: Schematic overview of the CMS L1 trigger system [120]

The L1 trigger is mostly comprised of high frequency, adjustable electronics that run over a minimal set of predefined instructions providing fast decisions within 4 μs of the collision. The L1 trigger system shown in Fig. 2.11 utilizes information from the calorimeters and the muon chambers, generating simpler physics objects called L1 candidates with the information of the energy clusters to make decisions. The L1 conditions are organized in a menu made of various algorithms, hard coded in the firmware of a Field-Programmable Gate Array (FPGA). Each sub detector provides Trigger Primitives (TP) which are further processed in many steps and combined in the Global Trigger (GT) to result in the final decision.

The second stage of the trigger decision falls upon the HLT consisting of an array of multi-core computers. The processing of the data at the HLT is structured around the concept of the HLT path, which is a set of algorithmic processing steps performed in

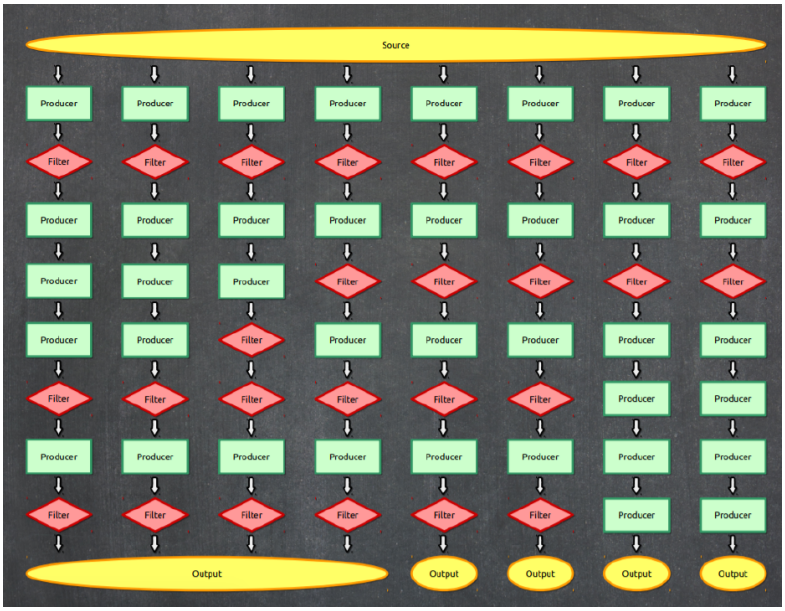


Figure 2.12: Schematic overview of the CMS HLT trigger system [135]

a predefined order as shown in the schematic Fig. 2.12. The selection of events is performed similar to offline processing of the "particle flow" (PF) algorithm discussed in Section 3.1. For events that pass the L1 trigger, the readout unit extracts all information from the CMS subdetectors which are built into detector segments, hits and clusters to form complete events. Lastly, a filter performs the reconstruction of the physics objects and determines the trigger decision.

Chapter 3

Object reconstruction

This chapter provides the details about the physics objects used in this analysis, namely photons and jets. They are necessary to perform the analysis motivated in Chapter 1. The CMS detector being a general-purpose detector utilizes a holistic approach to physics object reconstruction. The components of the CMS detector as described in the Section 2.2 have properties well suited to particle-flow (PF) reconstruction by taking advantage of an extensive list of the final-state particles identified and reconstructed to provide a global event description as described in Section 3.1.

The details about the photon reconstructed based on the PF reconstruction are provided in Section 3.2. This section provides details about the trigger system used to collect the data for this analysis and the reconstruction procedure for the photon. This is followed by the details of obtaining an identification criteria for the photon and corrections applied to the photons.

The studies of the substructure of jets is the primary goal of the analysis performed in this thesis. Section 3.3 provides the details about the jet objects with a brief discussion about various jet clustering algorithms focusing on the details about the reconstruction used in this analysis. Further details about the corrections applied to the jets and the tools developed to study the substructure of the jets are also provided in this section.

3.1 Particle flow reconstruction and global event description

As discussed in Chapter 2, the CMS detector is a very versatile detector consisting of various components making it applicable in a broad scope of physics goals. The presence of a highly segmented tracker, a fine grained electromagnetic calorimeter, a hermetic hadron calorimeter, a strong magnetic field and an excellent muon spectrometer makes the CMS detector well suited to particle flow reconstruction. The general goal of the detectors at hadron colliders is to reconstruct physics objects based on the signals collected by a given detector which can be summarized in a simplistic form as follows:

- The electrons and photons are primarily detected in the electromagnetic calorimeter.
- Jets primarily consist of hadrons and photons. Their energy can be measured by the electromagnetic and hadronic calorimeters.
- The tagging of jets originating from hadronic decays and heavy quarks pertains to the properties of charged particle tracks and in essence involves the tracker.
- The muon identification is achieved by considering the information from the dedicated muon detectors

However, considering these detectors in isolation poses limitations on the understanding of the physics objects. Thus, correlating the basic elements from all detectors to identify the final state particles allows for a holistic approach which is called the particle flow reconstruction [136]. The key ingredient to achieve this goal in hadron colliders is a fine spatial granularity as a coarser granularity may cause signals from different particles to merge. With sufficient granularity, a global event description becomes possible, in which all particles are identified. A schematic representation showing the PF reconstruction in the global description of objects used for analyses is shown in Fig 3.1.

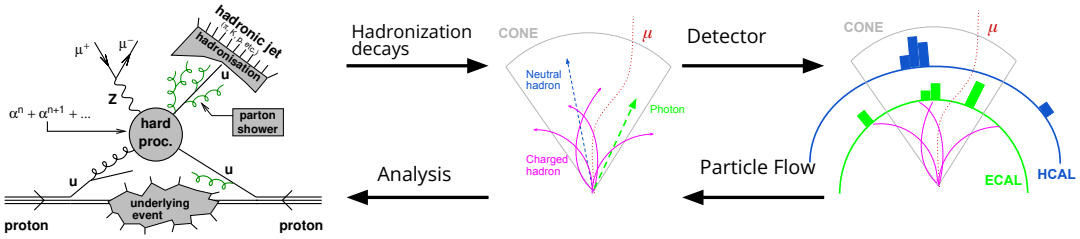


Figure 3.1: Schematic diagram of the steps involved in studying a physics process. A hard process such as parton scattering can occur at high energy collisions leading to the quark or gluon to hadronize. The final state decay products are observed in the CMS detector. PF reconstruction is the process through which the physics objects are reconstructed and further used in analysis [137]

The steps of the algorithm to reconstruct the basic PF elements in its essentials are, the reconstruction of the trajectories of charged particles in the inner tracker, with further segmentation into electron and muon track reconstruction; finally, the reconstruction and the calibration of calorimeter clusters. Charged particle tracks are reconstructed using an iterative tracking procedure. The first step is a combinatorial track finder based on Kalman Filtering [138] with the goal of measuring the momentum of energetic and individual muons, hadronic decays and tagging heavy quark jets. Stringent track quality criteria are implemented at this stage, primarily targeting energetic particles and well measured tracks to keep the rate of reconstruction of misidentified tracks at the level of a few per cent but limiting the efficiency. The hadrons traversing the trackers may also interact with the tracker material leading to kinks in the trajectory or production of secondary particles. By employing an iterative procedure for track identification with each iterative step having a different target such as prompt tracks at earlier steps followed by displaced tracks, then tracks inside jets and finally muon tagged tracks, each having a different but clear signature; the overall efficiency of the combinatorial track finder can be improved.

Electron reconstruction, aimed at characterizing energetic prompt electrons are primarily based on ECAL measurements utilizing a traditional seeding strategy called ECAL-based approach [139]. The position of the electrons are inferred using the cluster energy

and position. However most of the electrons emit bremsstrahlung photons carrying a significant portion of their energy. They are grouped together into a "supercluster" in a small window in η and an extended window in ϕ to account for the bending of the electron due to the magnetic field. For other electrons, such as electrons in jets and low p_{T} electrons, the energy is spread over an extended region which cannot be covered by the supercluster and with overlapping contributions from other particle deposits. Tracker-based electrons as part of the iterative tracking procedure in the PF reconstruction is designed to have a large efficiency for these electrons. A Gaussian-sum filter (GSF) [140] is more adapted to electrons than the Kalman Filter. The ECAL-based and tracker-based electrons are merged into a unique collection for the full electron tracking. Muon tracking is not specific to the PF reconstruction as the muon spectrometer allows muons to be identified with high efficiency and a high purity is granted by the calorimeters absorbing the other particles. The muon collections are composed for three different types: standalone muon, global muon and tracker muon.

The calorimeter clusters are another important component of the PF reconstruction with the purpose of detecting stable neutral particles such as photons and neutral hadrons, separating neutral hadrons from charged hadron deposits, reconstructing electrons and to provide additional information for the energy measurement of charged hadrons with inaccurate track parameters. A dedicated clustering algorithm was developed for PF event reconstruction to separate close energy deposits and detecting low energy particles by clustering separately in each subdetector. Crystals with energy larger than a threshold, called cluster seeds are taken to form topological clusters with its neighbors. Further reconstruction is performed by an expectation-maximation algorithm by considering a gaussian spread of energy in the neighboring crystals. Photons and neutral hadrons are reconstructed from these calorimeter clusters. Photons have a significant probability (expected to be around 60% from simulation) of converting into an electron-positron pair while traversing the material in front of the ECAL. A dedi-

cated conversion finder is utilized to avoid associating them with charged hadron and improve photon isolation determination [141]. Neutral particles do not leave any signature in the tracker and are thus identified by calorimeter clusters separated from any extrapolated charged particle track. Electrons and photons deposit their energy primarily in the ECAL while hadrons generally deposit energy in both ECAL and HCAL. The energy deposited in each of the detectors is lower than the true energy of the objects due to various factors such as transparency losses and the application of a minimum threshold to mitigate noise. The reconstruction involves a step to correct the detector losses by first applying a correction to the energy in the ECAL followed by the HCAL. The ECAL energy is corrected using simulations which are then verified with fits of photon pair to the π^0 invariant mass distribution in both simulation and data resulting in a per-cent agreement as shown in Fig.3.2 [122, 136]. Further corrections for the reconstructed objects are discussed in Section. 3.2.2. Following the corrections in the ECAL, the HCAL energy is calibrated using simulations of neutral hadrons leaving deposits in the HCAL.

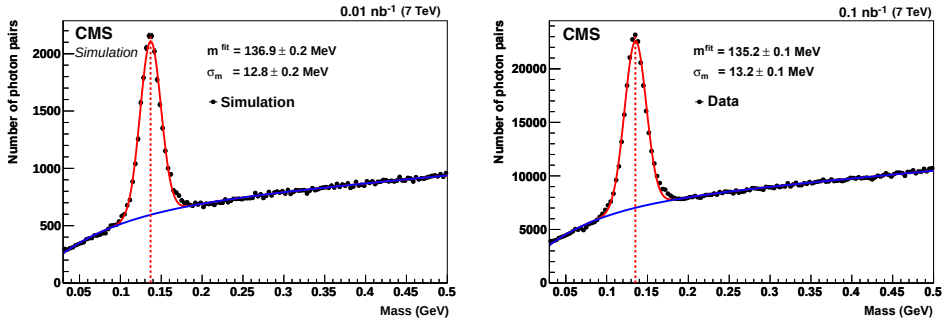


Figure 3.2: Photon pair invariant mass distribution in the ECAL barrel for simulation (left) and data (right). The π^0 signal is modelled by a Gaussian (red curve) and the background by an exponential function (blue curve) [136]

Physics objects, in general are composed of several PF elements described in the previous sections. The process of reconstructing a particle from the PF elements of different subdetectors proceeds with a link algorithm. The link algorithm tests parts of elements

in the events restricted to nearest neighbors in the (η, ϕ) plane to produce PF blocks of elements. The order in which the identification is performed is as follows:

- First, muon candidates are identified and the corresponding PF elements are removed from the block
- Then the electron identification is performed with the aim of collecting the energy of all bremsstrahlung photons along with prompt photons removing the associated tracks and clusters from further consideration
- Remaining tracks with a p_{T} uncertainty in excess of the calorimetric energy resolution for charged hadrons are masked.
- Charged hadrons, neutral hadrons and photons are associated to the available PF elements
- Secondary particles arising from hadronic decay are then identified and reconstructed.
- A final post processing step is implemented to reduce the particle misidentification and incorrect reconstruction.

Constituents are assigned to muons which are well described by the dedicated muon chambers and tracks using the tracker hits, followed by single electron and photons. Charged hadrons, neutral hadrons and photons arising from parton fragmentation, hadronization and hadronic decays are then assigned to jets. The calorimeters have a wide acceptance with the ECAL being up to $|\eta| < 3.0$ and the HCAL having forward detectors with an acceptance up to $|\eta| < 5.2$. The tracker acceptance region is more restrictive at $|\eta| < 2.5$ and determines the ability to distinguish charged and neutral hadrons as they are indistinguishable outside the acceptance. Thus, the procedure for the reconstruction of physics objects outside the tracker acceptance is different from those within the acceptance. Within the tracker acceptance, any ECAL clusters are pri-

oritized for photons over neutral hadrons, justified by the observation that for hadronic jets, 25% of the energy is carried by photons, while only 3% of the energy in the ECAL is due to neutral hadrons [136]. This justification does not apply in the case outside the tracker acceptance due to the presence of charged hadrons as well. Thus, beyond the tracker acceptance, ECAL clusters linked to HCAL clusters are assumed to be from the same hadron shower, while ECAL clusters without a linked HCAL cluster are considered photons. Since, the analysis presented in this thesis is limited to jets in the tracker acceptance region, further discussion of jets beyond the acceptance is not provided.

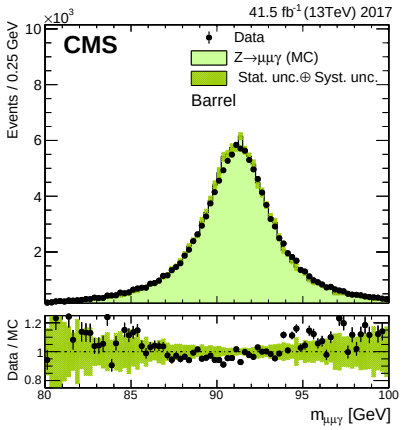


Figure 3.3: Invariant mass distributions of $Z \rightarrow \mu\mu\gamma$ shown for barrel in pp collisions comparing the data collected in 2017 with simulation[122]

The HCAL clusters are then linked to tracks, which might be linked to other ECAL clusters. The calibrated energy of the hadrons are obtained through the maximum of either the sum of the p_T of the tracks or the sum of ECAL and HCAL. Comparing the track momenta with the calibrated calorimetric energy provides the information about the particle content. If the calorimetric energy is compatible with the sum of the track p_T within the expected energy resolution, then the constituents are composed entirely of charged particles. If the calorimetric energy is larger, the excess is considered to be due to the presence of neutral hadrons and photons. Smaller excess energy in the order of 500 MeV in the ECAL are associated with photons while larger total calorimeter

energies are considered a combination of photons and neutral hadrons considering the proportion of energy in the ECAL and HCAL. Tracks are associated to charged hadrons with the track momentum used to assign the corresponding momentum and energy. In the case of large discrepancies between the calorimetric energy and track momentum, a relaxed search including muons is performed and included in the PF constituents for the jet. Thus, resulting PF constituents are classified under leptons, photons, charged hadrons and neutral hadrons which are clustered into jets and discussed in Section. 3.3.

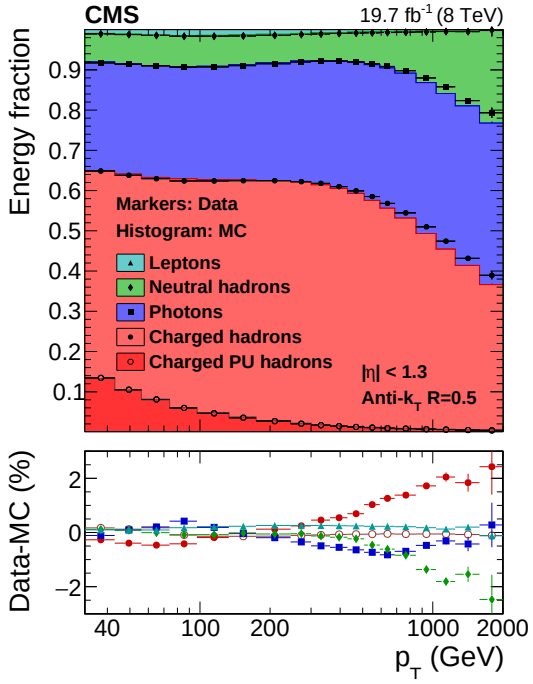


Figure 3.4: Jet energy composition in observed and simulated events as a function of p_{T}. The top panels show the measured and simulated energy fractions stacked, whereas the bottom panels show the difference between observed and simulated events[136]

This results in a global event description, when all the PF blocks have been processed and all particles have been identified. The resulting physics object performance is improved in all aspects; efficiency, purity, energy and angular resolution, thus, reducing the systematic biases in physics analyses. The performance of the PF reconstruction in pp collisions for photons is shown in Fig. 3.3. The energy scale difference between data and simulation, both from $\mathrm{Z} \rightarrow \mu\mu\gamma$ events, is smaller than 0.1% for photons[122].

The performance of the PF reconstruction in pp collisions for jets shown in Fig. 3.4 studied by comparing the energy of the PF constituents of the jets. The observed and simulated energy fractions agree within 1% for $p_{\mathrm{T}} < 500\,\mathrm{GeV}$, and within 2% for $p_{\mathrm{T}} \geq 500\,\mathrm{GeV}$ [136] .

3.2 Electron and photon

Photons and electrons are treated in a very similar manner in the CMS detector, with the significant difference being that electrons produce hits in the tracker layers. Both photons and electrons deposit almost all of their energy in the electromagnetic calorimeter (ECAL). Also, as photons or electrons propagate through the detector, it may interact with the material in front of the ECAL with photons converting into an electron-positron pair or electrons emitting bremsstrahlung photons. Thus, when the energy deposit is detected by the ECAL, the original single electron or photon may no longer be a single particle, but a shower of multiple photons and electrons.

The legacy reconstruction algorithm used for photon reconstruction in the CMS detector is known as the "Island algorithm" [124, 139]. Deposits of energy in the ECAL are clustered and an estimation of of the energy and position is made from this information. The Island algorithm starts by searching for seeds, which are crystals with an energy above a certain threshold as illustrated in Fig.3.5. The most energetic seed is considered the starting point, and the neighboring crystals are collected to form a cluster until it finds a rise in energy or a hole. Further superclusters are built to collect the energy of the photons decaying to electron-positron pairs (or the bremsstrahlung photon energy from electrons) by combining multiple clusters. While this simplistic approach has its merits and has been successfully used in past analyses, the CMS detector with it's highly versatile sub detectors can utilize the PF framework for physics object reconstruction as described in Section 3.1. The PF reconstruction of electrons and photons [122] has been developed and is used in pp collisions, but it's suitability to PbPb collisions was studied for the data collected in 2018 used in the analysis studied in this thesis. The global event description (GED) provided by the PF framework used in pp collisions is modified to be suitable for the large underlying event in PbPb collisions and is described in Section 3.2.2.

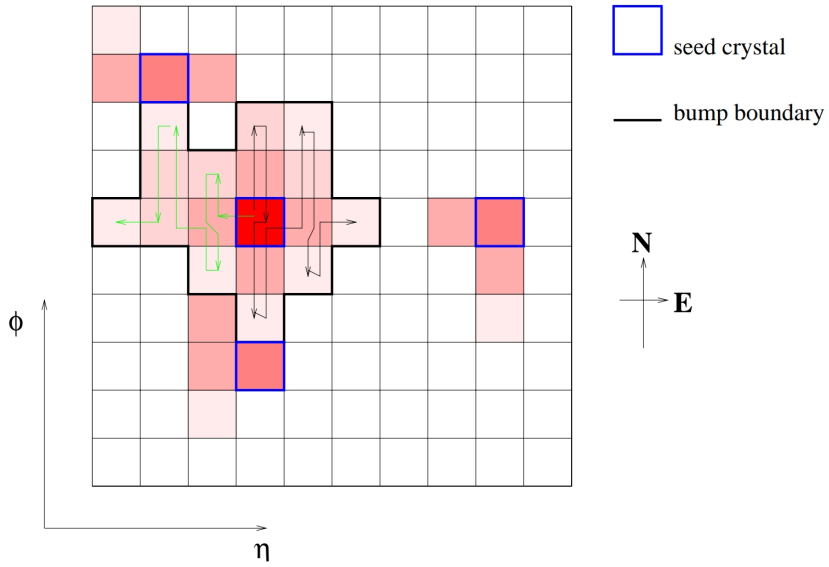


Figure 3.5: Illustration of the Island clustering algorithm in the ECAL[139]

The analysis presented in this thesis requires the presence of a high p_{T} photon which is used as a proxy for the initial unquenched jet p_{T} to facilitate a fair comparison between the PbPb and pp systems. Thus, the requirement of a photon is applied from the very first step in the event selection by requiring a photon triggered event. Further reconstruction and corrections are applied to this photon as discussed in the following sections.

3.2.1 Online reconstruction and triggers

The trigger system of the CMS is described in Section 2.2.7. The Level 1 (L1) trigger system, composed of hardware processors utilizes basic information from the individual subdetectors to construct L1 candidates. In the case of electrons and photons, the candidates are formed from trigger towers defined by arrays of 5×5 crystals in the ECAL barrel and a more complicated pattern in the endcaps. Due to the lack of tracker information at this stage, electrons and photons are not separated and considered as "EG" objects. A seed tower is identified as the crystal with the largest transverse energy above a fixed threshold ($E_T > 2\,\mathrm{GeV}$). Energy deposited at forward rapidities

in PbPb collisions are larger than in pp collisions. To suppress the large accept rates coming from forward rapidities an explicit rapidity selection was applied for EGs in PbPb collisions. The selection is $|ieta| \leq 24$, where *ieta* is a hardware–level rapidity and $|ieta| = 24$ corresponds to $|\eta| \approx 2.1$.

The events accepted by the L1 trigger are passed to a second trigger known as the High Level Trigger (HLT) which performs a simplified reconstruction of the physics objects performed offline. The use of GED reconstruction for both PbPb and pp collisions is implemented in a simplified form at the HLT as well. The details of the offline reconstruction are provided in the following section. The primary differences between online and offline reconstruction are driven by the limited CPU time available at HLT, the lack of final calibration and a more conservative selection criteria. The tracking algorithm for electrons at HLT is considerably different as the electrons are ECAL-driven, since tracker driven electron reconstruction is advantageous only for low energy or non-isolated electrons. The HLT electron and photon candidates are reconstructed from energy deposits in the ECAL crystals grouped into clusters around the corresponding L1 candidate. The ECAL clustering is performed only around the selected L1 candidate to form superclusters with electrons being associated with a track as well. Further identification and isolation criteria along with a minimum threshold on the E_T. The identification criteria are based on the amount of energy deposited in the HCAL downstream of the central tower, and on variables sensitive to the spatial extent of the electromagnetic shower [120]. The efficiency of the triggers were studied as a function of reconstructed photon energy as shown in Fig. 3.6 [121].

3.2.2 Offline reconstruction and corrections

The offline reconstruction of photons is performed using the global event description of the PF reconstruction[122]. Photons deposit almost all of their energy in the ECAL, without any hits in the tracker or energy deposits in the HCAL. Isolated photons typi-

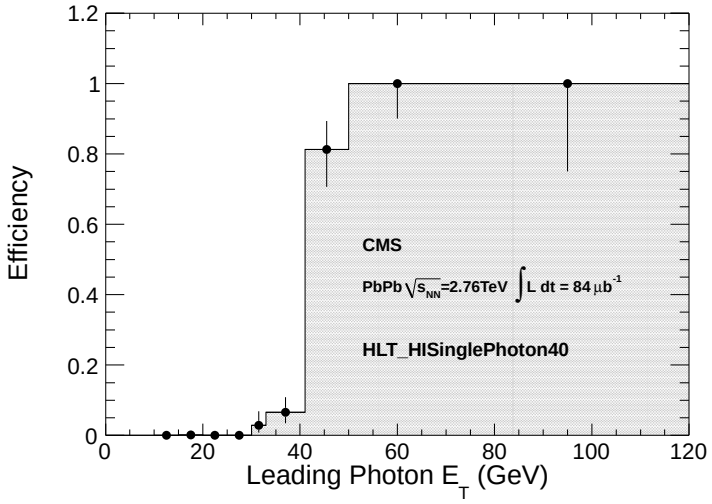


Figure 3.6: Trigger efficiency as a function of corrected offline photon transverse momentum in PbPb collisions at $\sqrt{s_{\mathrm{NN}}}$ =2.76 TeV [121].

cally interact with only one ECAL crystal but the deposited energy is spread over multiple ECAL crystals. Photons may interact with the nuclear matter in the material in front of the ECAL converting into an electron-positron pair. Thus, photons in general may not be a single particle when it reaches the ECAL, but a shower of multiple electrons and photons. A 3×3 matrix of ECAL crystals centered around the crystal with the highest energy contains about 95% of the energy of a photon, with a 5×5 matrix containing about 98%. The barrel (EB) and the endcap (EE) regions of the ECAL are generally treated independently during the reconstruction of photons due to differences in the structure of the crystals and the behavior of the photons. Since the analysis presented in this thesis utilizes pp and PbPb events with low pileup, out-of-time pileup subtraction is excluded from the reconstruction.

The general steps in the PF framework involved in the photon reconstruction are as follows:

1. The formation of clusters by grouping together crystals of energy is the first step. A predefined threshold of two to three times the order of electronic noise expected

is applied to these crystals (typically ≈ 80 MeV in EB and ≈ 300 MeV in EE). This results in a seed cluster, when a cluster is found to contain most of the energy deposited in a specific region with a minimum E_T above 1 GeV.

2. The second step is to combine clusters in a certain geometric area around the seed cluster into a supercluster (SC) to include photon conversions.

3. Electron reconstruction is performed in parallel with the information from the tracker as part of the iterative tracking described in Section 3.1

4. PF elements such as ECAL clusters, SCs and tracks along with conversion tracks are imported into consolidated PF blocks of particles using the link algorithm.

5. The PF blocks are resolved into photon objects starting from the SC or in the case of electrons from tracks. These PF electrons and photons are not differentiated in their treatment up to this step and the final linked ECAL clusters are promoted into a refined supercluster.

6. A loose selection requirement is applied to refined superclusters to distinguish electrons and photons. Objects passing the track selections are labelled as electrons and the failing objects are considered PF photons. These unbiased objects are used as the starting collection of PF photons and further corrections and selections are applied to them.

7. As a final step, stricter selection criteria are applied to separate electrons and photons from hadrons in the PF framework.

Electrons and photons can be composite objects consisting of several clusters and tracks and thus, present a difficult challenge in the PF reconstruction where hadronic constituents of jets may be misclassified as part of the electron or photon. A minimal selection criteria is thus required to correctly identify photons while improving jet and

missing transverse momentum measurements. A loose requirement on different BDT classifiers for isolated and nonisolated electrons are usually required, while photons candidates are required to be isolated with a shower shape expected of genuine photons. A dedicated conversion finder [141] is used to create links between any two tracks compatible with photon conversion and is treated differently than isolated photon candidates. The isolated photons obtained as a result of these selections are GED photons which are used in this analysis.

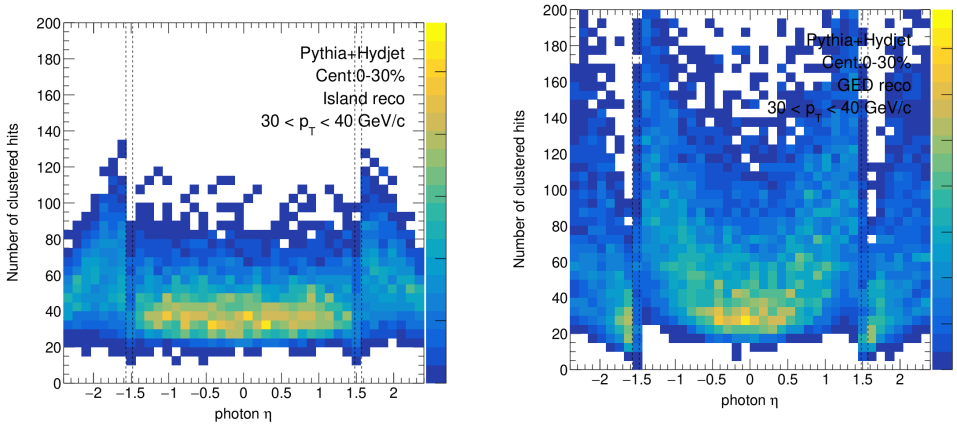


Figure 3.7: The number of clustered hits in simulated central PbPb events using the Island algorithm [139] (left) and the default PF algorithm [122] (right) adapted from the CMS internal analysis note AN-18-314 [142]

In the case of heavy ion collisions such as in the PbPb system, a deconfined state of matter knows as the quark-gluon plasma (QGP) is produced. Electroweak bosons such as photons and Z bosons that decay into leptons do not interact through with the QGP through the strong interaction and thus, the p_{T} of the boson reflects the initial energy of the associated parton before any energy loss. Modifications are made to the standard photon and electron reconstruction to account for the high particle multiplicity in central PbPb collisions. The distribution of the number of clusters in simulation are compared between the previously used Island algorithm and the default GED algorithm in Fig. 3.7. There is a significant increase in the number of clusters formed by the GED algorithm for central PbPb events. Thus, a stricter requirement is placed

while constructing superclusters by combining clusters geometrically around a seed cluster. Generally, a dynamic window is chosen such that the energy spread following a Gaussian profile can be recovered using the assumption that local maxima above 1 GeV threshold correspond to a single particle incident. Due to the magnetic field in the CMS detector, the distribution of energy has a bend in the ϕ direction. In the case of PbPb events, which have a denser environment of underlying events, this algorithm gives worse energy resolution and misidentification rate. A fixed-width approach was applied where the size of the SC was restricted to 0.2 in the ϕ direction, resulting in the misidentification rate decreasing from 2.7% to 0.5% and a energy resolution under 8% (3%) for photons above 20 GeV (100 GeV) as shown in Fig.3.8. The superclusters were also required to have at least 15 GeV in an effort to reduce the reconstruction time.

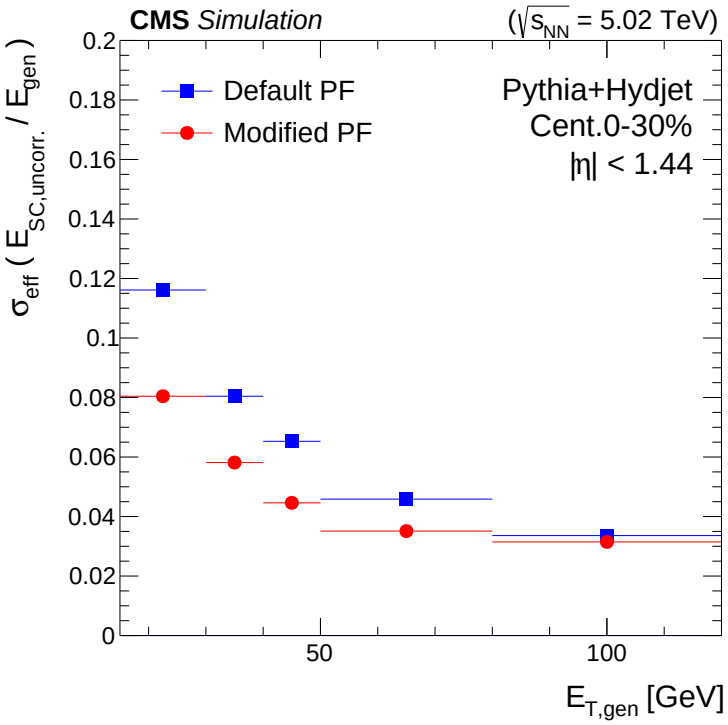


Figure 3.8: Comparison of the photon energy resolution in simulation for the default and modified PF reconstruction in PbPb collisions as a function of the true photon E_T[122]

The reconstruction of the photon does not necessarily recover a photon's energy ac-

curately as shown in Fig.3.9. Photons reconstructed as superclusters in the ECAL is subject to energy loss through multiple sources such as leakage, intermodule gaps or energy lost in the tracker. In the case of pp events, the energy of reconstructed photons in simulation was found to be recovered well and no corrections were applied. In the case of PbPb events there is a noticeable discrepancy which is corrected using a regression analysis [142].

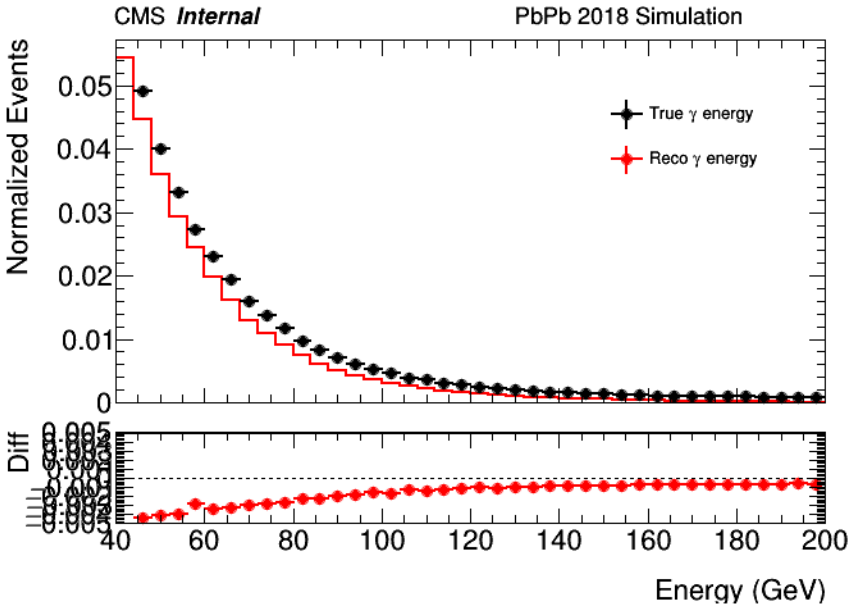


Figure 3.9: Comparison of the true and reconstructed photon energy in simulation for PbPb collisions as a function of the true photon E_T

A multivariate regression using the Toolkit for Multivariate Data Analysis [143] is used to correct the energy estimation and improve the energy resolution. The regression technique described here uses simulation events only, with a requirement of generator level photons matched to reconstructed photons. The target in the training is the energy of the gen-level photon with the information about the SC and underlying event (UE) activity used as input for the training as described below:

- The raw SC energy, η and ϕ are provided as the basic quantities identifying the photon.

- The crystals within the SC are weighted by their energy to calculate the width of the SC in the η and ϕ directions.
- The crystals with the highest, second highest and the total energy in the 3×3 matrix of crystals are provided to distinguish photon decays by the spread outside of single crystal signatures.
- The energy distribution within the supercluster is the most important information to recover the photon's energy. Asymmetries of the energy distributions in the η and ϕ directions are combined with respect to the maximum energy crystal and provided as an input. At second order, covariances of energy in a 5×5 matrix of crystals are provided
- Finally, to account for the UE, the median transverse energy density of the event is considered.

The energy scale and resolution of the photon's energy after applying the corrections are shown in Fig.3.10 with the gaussian and double sided crystal ball (DSCB) fits of the ratio of the reconstructed and true p_{T} [142]. The energy resolution of the photon after the corrections is of order 2%.

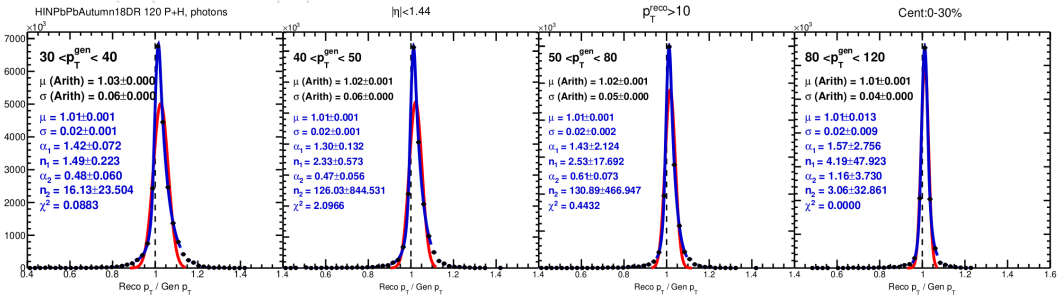
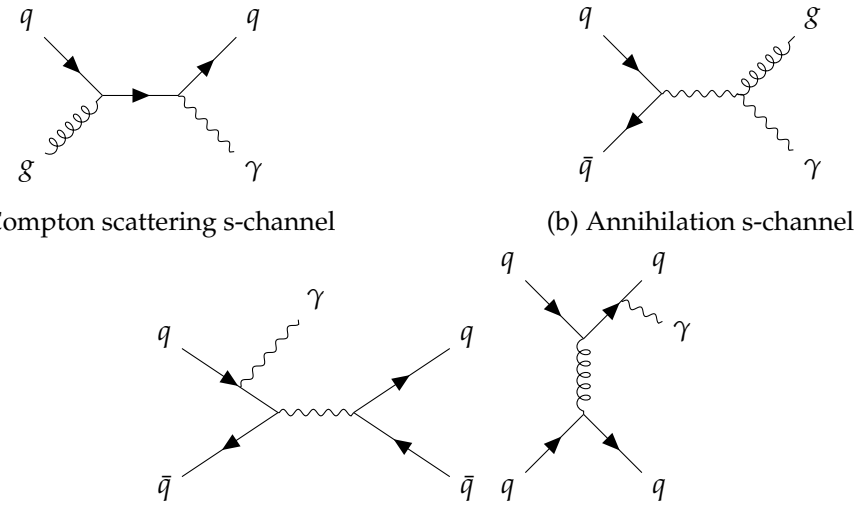


Figure 3.10: The reconstructed p_{T} over true gen p_{T} distributions for photons in simulated PbPb events fit with a Gaussian(red) and DSCB (blue) functions in bins of true gen p_{T} [142]

3.2.3 Prompt photon identification

Photon signal and background



(a) Compton scattering s-channel

(b) Annihilation s-channel

(c) Examples of some NLO processes

Figure 3.11: Feynman diagrams for the various processes producing a prompt photon

Photons being detected by the CMS detector are produced through many different channels. However, the QCD processes containing photons of interest generally require the photon to be from the hard scattering of the partons in the hadronic collision. We consider these photons to be "prompt" photons. Prompt photons are produced at leading order in pertubative QCD through one of 3 processes shown in Fig.3.11. They are (a) quark–gluon Compton scattering, (b) quark–antiquark annihilation, and (c) collinear fragmentation of a final-state parton into a photon. Photons from Compton scattering and annihilation processes are considered "direct" photons, while photons from parton fragmentation are called "fragmentation" photons in this context. The production cross-section of photons has been studied in vacuum through pp collision systems as shown in Fig.3.12 [144–146]. Recent measurements for both low and high p_{T} isolated photons with a strong agreement to perturbative QCD predictions are shown in Fig.3.13 [147, 148].

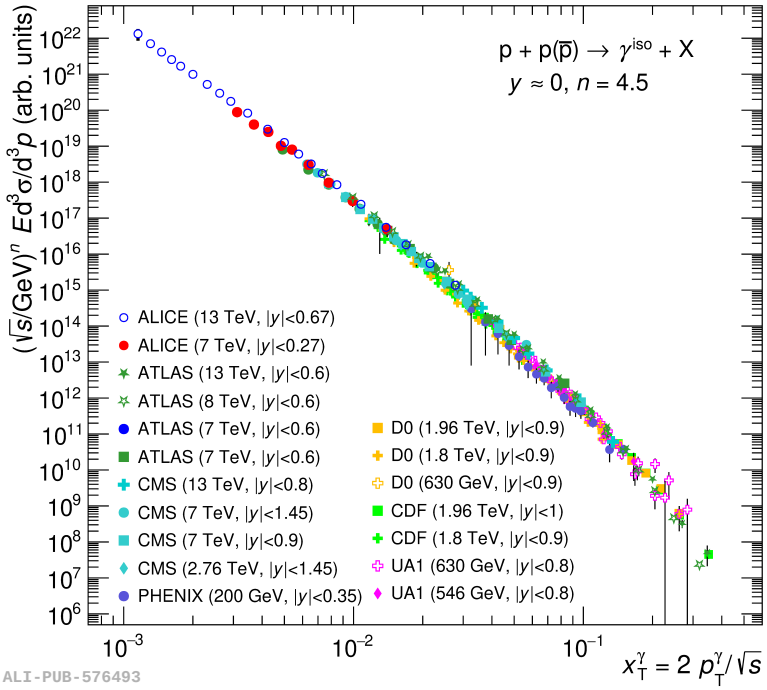


Figure 3.12: Comparison of isolated photon spectra measured in pp (p $\overline{\text{p}}$) collisions by various experiments compiled in Ref.[144] and adapted from Ref.[146]

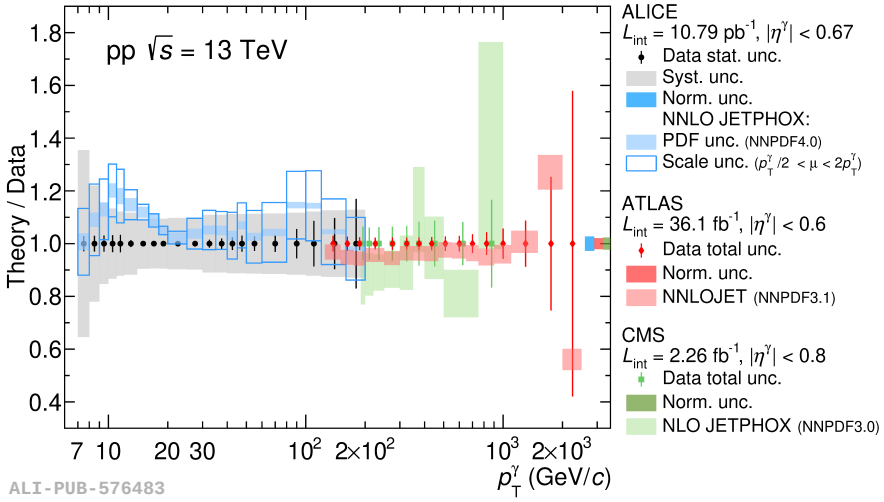


Figure 3.13: Isolated photon cross section compared to theory measurements by ALICE, ATLAS and CMS for various p_{T} regimes [146–149]

Prompt photons are also measured in heavy ion collisions and are consistent to the vacuum measurements [100–104]. Thus, they are a clean probe of the initial state of the hadronic collisions as discussed in Chapter 1. However, the measurement of prompt photon production is complicated by the presence of a large background coming from the electromagnetic decays of neutral mesons such π^0 or η as shown in Fig. 3.14. Non-prompt photons originate from these hadron decays as the dominant decay mode for the pion is into 2 photons with a branching ratio of 0.98823. The production cross-section of these background dominated photons are shown in Fig.3.15 in pp and PbPb collisions. Neutral mesons can be highly boosted, particularly if they are produced in the fragmentation of high p_{T} jets. In that case, the neutral meson decay daughters (two photons in this case) will be highly boosted. Due to the finite granularity of the electromagnetic calorimeter of CMS, the two-photon decay of neutral mesons can be misidentified as prompt isolated photons. This is addressed in the analysis via background subtraction as discussed in the following section.

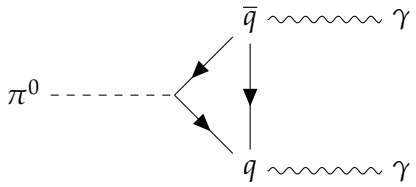


Figure 3.14: Feynman diagram for the decay of a neutral pion to a photon pair

For parton energy loss studies in heavy-ion collisions, a selection of direct photons is desired because they balance precisely the recoiling parton energy. The comparison of the contribution of fragmentation and direct photons to the production of isolated photon events is shown using LO simulations in PYTHIA8 [75] in Fig.3.16. According to simulations using LO matrix elements for the hard scattering, the sample of photons used in this analysis is dominated by prompt photons from quark-gluon Compton scattering and quark-antiquark annihilation processes. An alternative NLO simulation using JETPHOX [144] coupled to the fragmentation function of parton to photon is

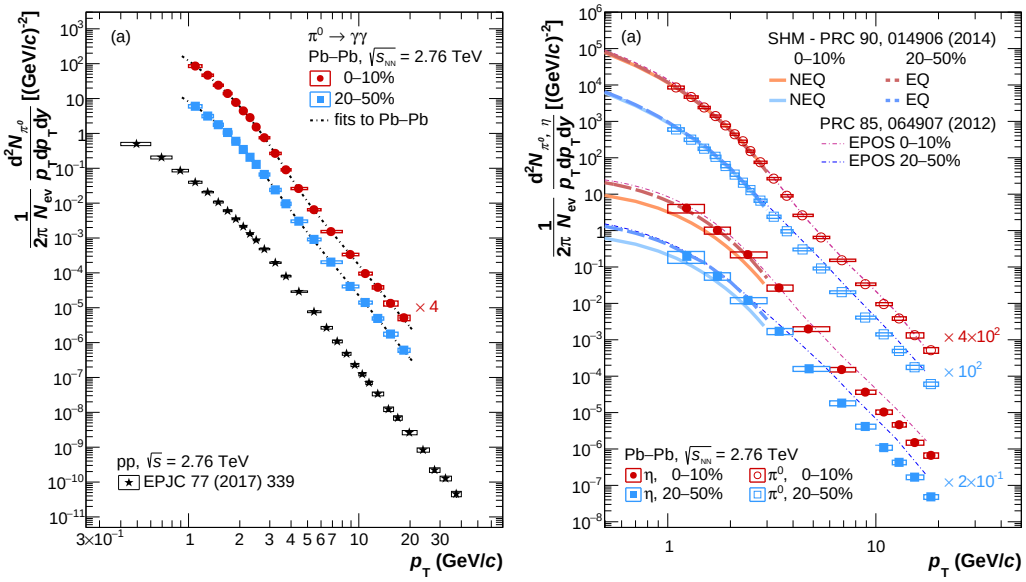


Figure 3.15: Cross section of π^0 decay to a photon pair in PbPb data compared to pp data (left) and to model predictions (right) [150]

also shown in Fig.3.16. The contribution of fragmentation of photons is of the order of 15% at high p_T. In order to suppress fragmentation photons, we apply isolation selection event requirements that are based on the fact that a fragmentation photon is generally surrounded by other products of the parton fragmentation. The procedure followed to obtain the selection criteria and background subtraction are described in the following sections.

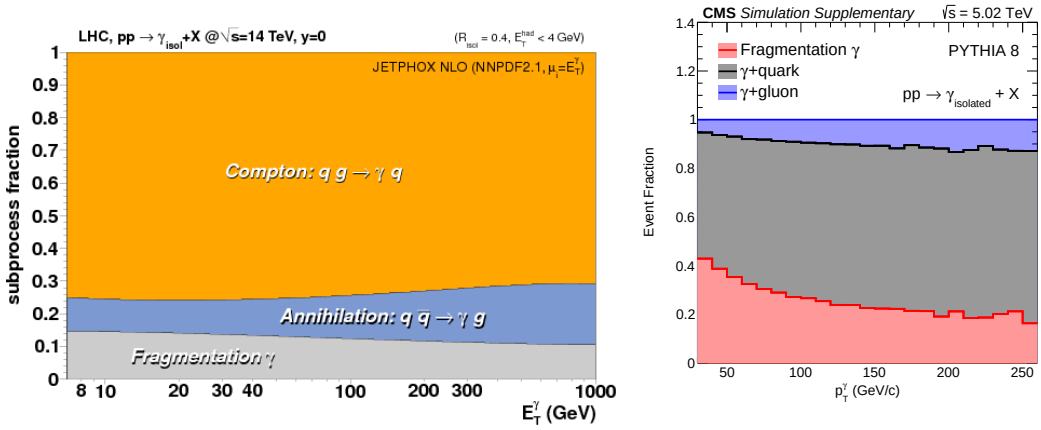


Figure 3.16: Fractional contribution of direct and fragmentation photons in simulations using JETPHOX (left) [144] and PYTHIA8 (right) [75]

Photon selection and discriminating variables

A common technique used in CMS for the identification of photons is based on sequential requirements (cut-based) of various discriminating variables to separate signal from background sources. Discriminating variables are generally classified as shower shape criteria and isolation criteria. Shower shape variables exploit the shape of the electromagnetic shower in the ECAL which can help identify hadron decays, which have a characteristic wider shower profile even if the photons themselves cannot be resolved. Isolation criteria considers sums of reconstructed energies around the photon which are a characteristic of particles resulting from parton fragmentation and in jets.

Shower shapes consist of variables constructed from the electromagnetic shower profile in the calorimeters. They are used to reject jets with high electromagnetic content and can identify hadron decays. Neutral hadron decay showers, which is the primary background source for photons, have a characteristic wider profile compared to prompt photons due to the pair production even when the individual photons cannot be resolved. Shower profiles along the ϕ direction are less discriminating than the η direction due to the effect of the magnetic field, which elongates electromagnetic clusters in the ϕ direction. Several shower shape variables are constructed to parameterize the difference between the shape of energy deposits from photons and hadrons from jets. The most relevant variables used for the photon identification are described below.

Fraction of hadronic energy (H/E): ratio of the HCAL energy over the photon supercluster energy. There are three sources that significantly contribute to the measured hadronic energy (H) for a prompt photon: leakage of electrons and photons through intermodule gaps, pileup and HCAL noise. Due to larger energy deposits on average in PbPb collisions as well as low pileup, these genuine contributions are negligible in our consideration. Since most of the background photons are produced inside jets, the amount of hadronic energy around the photon candidate is significantly larger than for

signal photons. This is quantified by using the ratio of HCAL over ECAL energy inside a cone of radius $\Delta_R = 0.15$ around the photon candidates.

Energy weighted covariance in η($\sigma_{\eta\eta}$): a measure of how the energy is spread in the ECAL. This is quantified via the shower-shape variable $\sigma_{\eta\eta}$, defined as

$$\sigma^2_{\eta\eta} = \frac{\sum_i^{5\times5} w_i(\eta_i - \eta_{5\times5})^2}{\sum_i^{5\times5} w_i}, w_i = max(0, c + ln\frac{E_i}{E_{5\times5}}), c = 4.7 \tag{3.1}$$

In further discussion, this quantity is simply called shower shape. The weight factor w_i is selected such that energies below the ECAL noise threshold are rejected and deposits of at least 0.9% of $E_{5\times5}$ will contribute to the shower shape. This variable has a strong dependence on the distance between two crystals in η. The size of the crystals in η in the ECAL barrel is 0.0175, while the endcap varies from 0.0175 to 0.05. Thus, the spread of $\sigma_{\eta\eta}$ in the endcap is expected to be twice that of the barrel. Background photons from neutral meson decays tend to have a wider spread of energy in the cluster while single photon has a narrow shower.

Isolation($\mathcal{I}$) is a generic class of discriminating variables that are constructed from the sum of reconstructed energies in a cone around the photon in different subdetectors. The cone is defined with a radius ΔR in $\eta - \phi$ centered at the supercluster position of the photon. To exclude the energy of the photon itself, a veto region inside the isolation cone needs to be defined. Background photons tend to have larger isolation energy around them, as a larger fraction of them are produced inside jets.

There can be particles that fall inside the isolation cone, but are uncorrelated to the nucleon-nucleon interaction that produce the photon. In pp events, these particles can originate from different pp interactions in the same bunch crossing (pileup). In PbPb events, they can originate from different nucleon-nucleon interactions in the same PbPb collision. It is hard to identify the particles that are uncorrelated to the interaction that

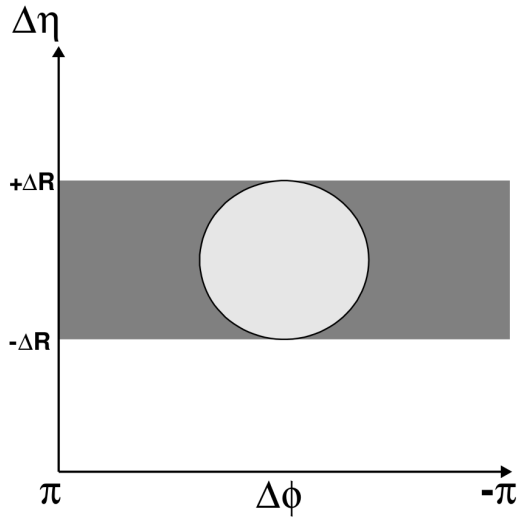


Figure 3.17: Schematic representation of the region considered for the normalization scheme in the UE subtraction in photon isolation calculation[142].

produce the photon object of interest. Therefore, contribution of the uncorrelated particles to an isolation variable must be estimated and subtracted. The isolation is calculated independently for the 3 sub detectors: ECAL isolation ($\mathcal{I}_{\text{ECAL}}$), HCAL isolation ($\mathcal{I}_{\text{HCAL}}$) and tracker isolation ($\mathcal{I}_{\text{TRK}}$). The sum of the 3 isolation energies after UE subtraction is considered the final isolation variable ($\mathcal{I}$) used in further analysis. The UE subtraction procedure for the photon isolation is performed by considering a region $\eta \in (\eta^{\text{SC}} - \Delta R, \eta^{\text{SC}} + \Delta R)$ and $\phi \in (0, 2\pi)$ as shown in Fig. 3.17[100, 142]. The isolation region is defined as a larger cone around the centroid of the photon supercluster in the ECAL, HCAL and tracker with the background estimated using the uncorrelated energy deposit in the ϕ direction, depicted by the rectangular region in the figure excluding the circular isolation cone itself.

The UE-subtracted detector based isolation variables for a cone of radius $\Delta R = i/10$ are defined as,

$$\mathcal{I}_i^{\text{DET,UEsub}} = \mathcal{I}_i^{\text{DET}} - \mathcal{I}_i^{\text{DET,UE}} \qquad \text{DET} = \text{ECAL, HCAL, TRK} \tag{3.2}$$

where the $\mathcal{I}_i^{\mathrm{DET}}$ denotes the energy inside the isolation cone while $\mathcal{I}_i^{\mathrm{DET,UE}}$ denotes the normalized energy estimated using the UE strip as follows,

$$\mathcal{I}_i^{\mathrm{ECAL}} = -E_{\mathrm{T}}^{\mathrm{SC}} + \sum_{\Delta R < i/10} E_{\mathrm{T}}^{\mathrm{ECAL}} \qquad \mathcal{I}_i^{\mathrm{ECAL,UE}} = c_R \sum_{(\eta,\phi)\in \mathrm{UEstrip}} E_{\mathrm{T}}^{\mathrm{ECAL}} \tag{3.3}$$

$$\mathcal{I}_i^{\mathrm{HCAL}} = \sum_{\Delta R < i/10} E_{\mathrm{T}}^{\mathrm{HCAL}} \qquad \mathcal{I}_i^{\mathrm{HCAL,UE}} = c_R \sum_{(\eta,\phi)\in \mathrm{UEstrip}} E_{\mathrm{T}}^{\mathrm{HCAL}} \tag{3.4}$$

$$\mathcal{I}_i^{\mathrm{TRK}} = \sum_{\Delta R < i/10} p_{\mathrm{T}}^{\mathrm{TRK}} \qquad \mathcal{I}_i^{\mathrm{TRK,UE}} = c_R \sum_{(\eta,\phi)\in \mathrm{UEstrip}} p_{\mathrm{T}}^{\mathrm{TRK}} \tag{3.5}$$

where the "UE strip" is defined by $(\eta^{\mathrm{SC}} - \Delta R, \eta^{\mathrm{SC}} + \Delta R), \phi \in (0, 2\pi)$ and $\Delta R > i/10$. The c_R factor scales the energy density in the UE strip to the isolation cone due to the different cross sectional areas of the 2 regions as,

$$c_R = \frac{\pi R^2}{2\Delta R \cdot 2\pi - \pi \Delta R^2} \tag{3.6}$$

$$= \frac{\Delta R}{4 - \Delta R} \tag{3.7}$$

resulting in the final summed isolation variable $\mathcal{I}$ defined as,

$$\mathcal{I} = \mathcal{I}_i^{\mathrm{ECAL,UEsub}} + \mathcal{I}_i^{\mathrm{HCAL,UEsub}} + \mathcal{I}_i^{\mathrm{TRK,UEsub}} \tag{3.8}$$

The distributions for the H/E, $\sigma_{\eta\eta}$ and $\mathcal{I}$ in simulated signal and background PbPb events are shown in Fig. 3.18. Since the isolation variable consists of an averaged energy sum of the UE, an over correction can lead to negative isolation values. This is corrected through the unfolding procedure as discussed in Section 4.7.

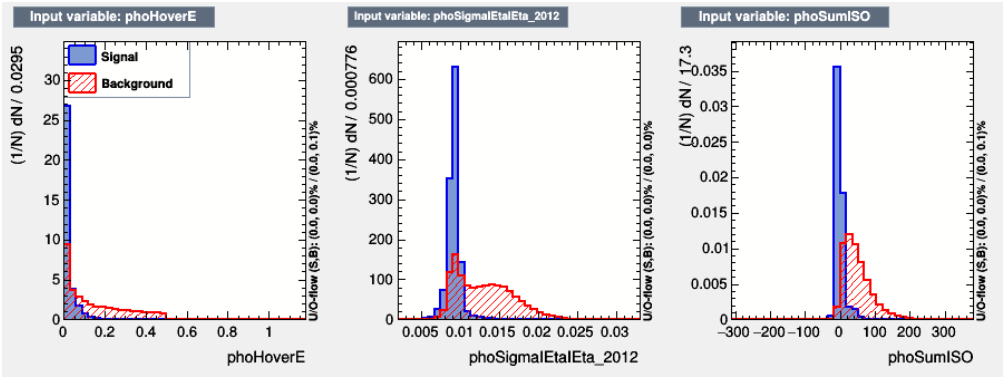


Figure 3.18: An example plot of the distributions of the variables H/E, $\sigma_{\eta\eta}$ and $\mathcal{I}$ showing the separation of signal (blue) and background (red) in simulated PbPb events.

Photon identification training and performance

For parton energy loss studies in heavy ion collisions, a selection of prompt photons is desired since they provide a more reliable proxy for the momentum of the recoil parton that initiated the jet shower. To suppress the contribution from nonprompt photons, a cut-based approach is followed using the genetic algorithm to optimize the search of approximate cut values based on the efficiency [143]. The signals at the particle level are prompt photons with isolation criteria such that the scalar p_{T} sum of all the particles around the photon within a radius of 0.4 in units of η–ϕ must be less than 5 GeV. Residual inefficiencies for the signal are accounted for in the corrections described in Section 4.7. The detector-level variables used in the optimization of the photon selection are the fraction of hadronic energy around the photon candidate H/E, the shower shape variable $\sigma_{\eta\eta}$, and the isolation variable $\mathcal{I}$ [122] as discussed in the previous section. The fraction H/E is calculated as the ratio of HCAL over ECAL energy within an η–ϕ distance of 0.15 units with respect to the photon candidate. The $\sigma_{\eta\eta}$ variable quantifies the lateral energy spread in the ECAL cluster, which is typically wider for photons from neutral-hadron decays than for prompt photons. The isolation variable $\mathcal{I}$ is given by the sum of transverse energies in ECAL and HCAL and the transverse momenta of all tracks with $p_{\mathrm{T}} > 2$ GeV within an η–ϕ distance of 0.4 with respect to the

photon candidate, which is corrected by subtracting the estimated average energy from the UE [100]. The $\mathcal{I}$ variable can have negative values as a result of the aforementioned UE subtraction. In PbPb (pp) events, we select isolated photons with $\mathcal{I} < 2.1\,\text{GeV}$ ($-0.1\,\text{GeV}$), which corresponds to a particle level $\mathcal{I}$ value of less than 5 GeV for generated photons. The distributions for these observables in the simulated samples used to obtain the identification working points are shown in Fig. 3.18. The performance of the training is evaluated through a background rejection vs signal efficiency curve as shown in Fig. 3.19 in the case of PbPb [95]. Tables 3.1 and 3.2 list the cut values derived for the discriminating variables for the specified working points for 2018 PbPb and 2017 pp respectively [142].

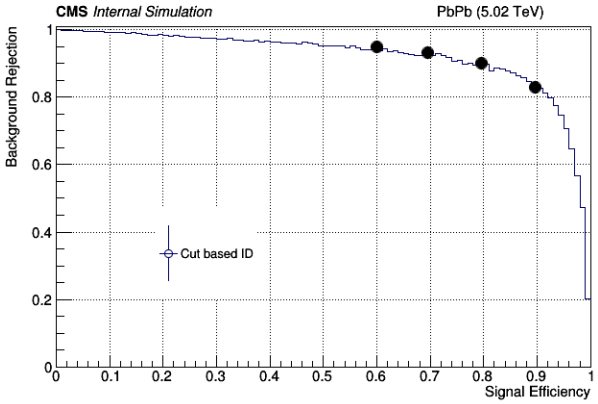


Figure 3.19: Background rejection vs signal efficiency curve (ROC) curve obtained in the TMVA analysis using the set of separation variables as discussed following the procedure used in the AN-18-314(Internal) [142]. The black circles denote the chosen working points at a signal efficiency of the training MC dataset of 60%, 70%, 80% and90% defined as extra-tight, tight, medium and loose working points

Table 3.1: Cut based ID obtained using simulated signal and background samples in TMVA for 2018 PbPb for the 3 photon identification variables [142]

	Loose (90%)	Medium (80%)	Tight (70%)	Extra-Tight (60%)
H/E	0.247995	0.238094	0.164101	0.119947
$\sigma_{\eta\eta}$	0.012186	0.011024	0.010784	0.010392
$\mathcal{I}$	11.697505	6.787425	3.509457	2.099277

Table 3.2: Cut based ID obtained using simulated signal and background samples in TMVA for 2017 pp for the 3 photon identification variables [142]

	Loose (90%)	Medium (80%)	Tight (70%)	Extra-Tight (60%)
H/E	0.081456	0.063521	0.009763	0.009732
$\sigma_{\eta\eta}$	0.011012	0.010886	0.009906	0.009905
$\mathcal{I}$	3.246891	1.549756	0.663719	-0.014755

3.2.4 Photon purity estimation

The contribution of photon-tagged jets originating from the decays of neutral hadrons is taken into account via signal photon purity of the cut-based selected sample. The photon purity estimation is obtained by performing a template fit of the shower shape distribution where the signal template is MC-based and the background template is built with nonisolated ($10 < \mathcal{I} < 20\,\text{GeV}$) photons in the data [100]. The region $10 < \mathcal{I} < 20\,\text{GeV}$ is depleted in signal contributions from prompt photons according to simulation studies. An alternative matrix ("ABCD") method is used to estimate the photon purity, which relies on data and is independent of the template fit method. The ABCD method consists of dividing the photon+jet data sample into four mutually exclusive regions using a two-dimensional plane with the variables $\sigma_{\eta\eta}$ and $\mathcal{I}$. We assume that the selection efficiencies of $\sigma_{\eta\eta}$ are independent of $\mathcal{I}$ for the background, which is supported by simulation studies. Three of the four regions are dominated by the background (the B, C and D regions), while the fourth region (A) is a mixture of signal and background. The regions A, B, C, and D are determined using the same selection requirements on $\sigma_{\eta\eta}$ and $\mathcal{I}$ as for the template fits. According to simulation studies, it is expected that the amount of remaining prompt photon events in the B, C, D regions is negligible with respect to the background. The signal inefficiency from the ABCD or template-fit methods, which is at the per mille-level, is taken into account as part of the corrections described in Section 4.7. The normalization of the background calculated with the ABCD method is different from the one obtained with the template fit by a few percent for PbPb and pp. The estimated photon purities have values of

0.77 ± 0.01 and 0.93 ± 0.02 for PbPb and pp collisions, respectively, where the central value corresponds to the photon purity obtained with the template fit method and the uncertainties correspond to the symmetrized difference of the photon purity obtained with the alternative ABCD method as shown in Figs. 3.20 and 3.21.

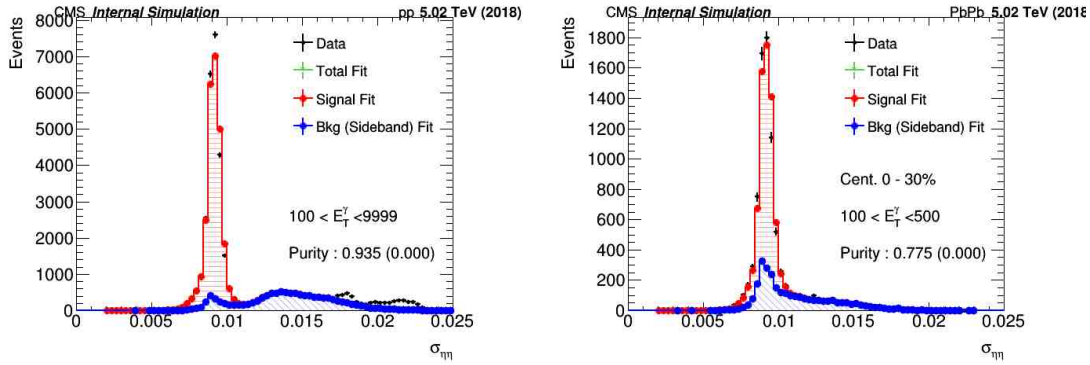


Figure 3.20: Extraction of photon purity using template fits for pp (left) and 0-30% central PbPb (right). The red histogram represents the signal distribution from MC (prompt photon production) and the blue histogram represents the distribution obtained from the sideband region from data, which is enriched in $\pi^0 \rightarrow \gamma\gamma$ decays.

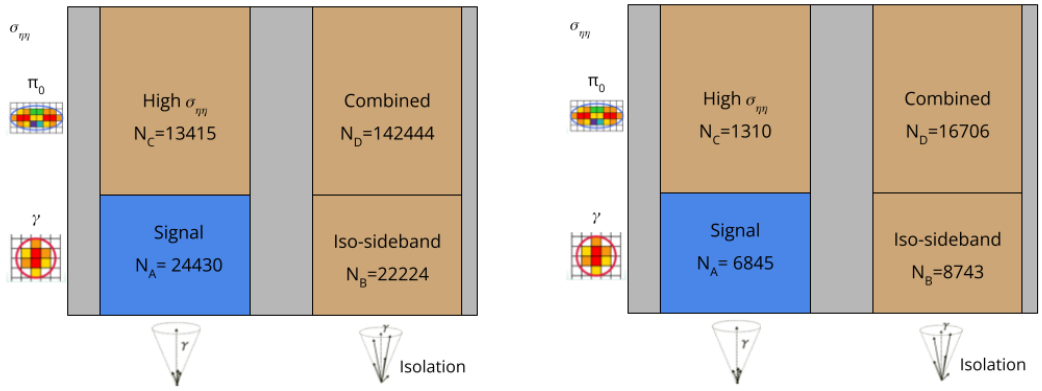


Figure 3.21: Extraction of photon purity using ABCD method for pp (left) and 0-30% central PbPb (right). The ABCD regions considered for the method are labelled in the plots. The region "A" corresponds to the target signal region for prompt photons.

3.3 Jet

Jets in the CMS detector are composite objects reconstructed by using a combination of multiple detector subsystems. As discussed in Chapter 1, jets are formed as part of the process of hadronization and the decays of these hadrons. The composition of jets shown in Fig. 3.4 illustrates the various components that can constitute the jet, primarily charged hadrons along with photons and neutral hadrons. Leptons are also a small component of the jet energy. Thus, the hadronic energy detected in both the calorimeters with the photons primarily depositing their energy in the ECAL, and with the identification of charged tracks in the tracker, all play a part in the jet reconstruction. In the following sections, we discuss the PF reconstruction procedure followed for jets in pp and PbPb collision systems. A further discussion on the energy corrections for jets is provided with an emphasis on background subtraction relevant to PbPb collisions due to the underlying event.

3.3.1 Jet reconstruction

The reconstruction of jets, similar to photons discussed previously is part of the PF framework which is a holistic reconstruction of the entire event discussed in Section 3.1. The PF constituents of the jet mainly consist of charged hadrons, followed by photons and neutral hadrons as shown in Fig. 3.4. The PF constituents for the jet are clustered using the clustering algorithms discussed in Chapter 1 such as the k_{T}, anti-k_{T} or CA algorithms using the FastJet framework [151]. The most commonly used jet clustering algorithm is the anti-k_{T} algorithm, due to its resilience to back-reaction and the almost circular shape of the jet area. Jets are clustered from the list of PF candidates using the anti-k_{T} algorithm [18] with a specified distance parameter which is a proxy for a circular jet radius of R in the y–ϕ plane. The four-momentum of the jet is determined using the vector sum of all particle momenta in the jet known as the E-scheme. Typically jets are reconstructed for $R = 0.2, 0.3, 0.4$ such that they are fully contained within the tracker

acceptance. The radius of the jet affects the background contamination within the jet cone and the different radii are used to study different physics effects. Jet reconstruction performance studies are required to discriminate between light quark and gluon, top quarks, W boson as well as jets from noise and pileup topologies [152]. Figure 3.22 shows the discrimination of quark vs gluon jets in pp collisions in Z +jets and dijets. The Z +jets are dominated by quarks while the dijets are dominated by gluons and shows that the MC models describe the data.

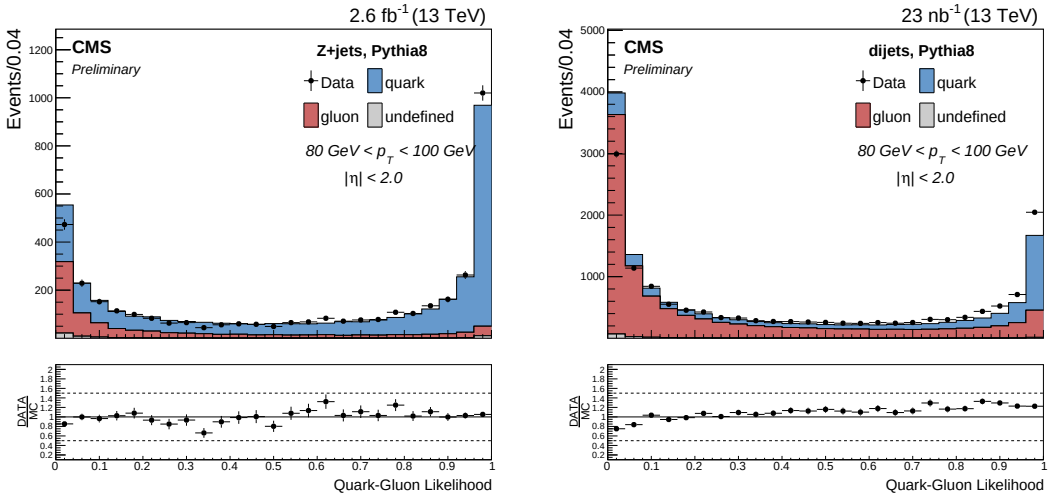


Figure 3.22: Example of a performance plot for quark-gluon jet discrimination in pp collisions in Z +jets(left) and dijet (right) [152]

In the case of heavy ion collisions, the PF reconstruction used in pp collisions needs to be modified due to the high track multiplicity [153]. This primarily affects the charged hadron constituents of jets and an alternative approach is used in their reconstruction for jets. A requirement is placed on the tracks such that all three layers of the pixel detector have a hit to reduce the fake rate. This reduces the efficiency to 60-70% compared to the expected 90% in pp collisions. The procedure followed for jet reconstruction is to first subtract the underlying event in the PbPb collisions. The procedure followed for the UE subtraction is discussed in Section 3.3.1. The UE background has regional fluctuations as well, which are further corrected through unfolding of the physics ob-

servables as discussed in 4.7. The jet energy corrections are derived following the procedure for pp collisions, with an additional correction applied using simulation for the UE background. The procedure for obtaining the jet calibrations is described in Section 3.3.2.

Underlying event background subtraction

The UE background subtraction is an important aspect of the jet reconstruction in PbPb collisions. The UE in central PbPb collisions is dependent on the jet area and p_{T} cuts on the tracks considered. Generally for jets of radii $R = 0.3$, the UE can be of order 40 GeV [154] with azimuthal modulations due to background fluctuations. These fluctuations are due to the overlap between the ions as discussed in Section 1.3.3 leading to an elliptic shape of the medium. This uncorrelated background needs to be subtracted from the jet constituents for a robust reconstruction of jets in heavy ion collisions [155, 156]. An event wise average background density (ρ) is estimated through an iterative pedestal subtraction technique [157] which is used for the background subtraction. The background contamination is estimated in strips of η to account of the geometry in the detector through the transverse momentum density ρ given by,

$$\rho = \mathrm{median}\{\frac{p_{T,i}}{A_i}\} \tag{3.9}$$

where i indicates the cluster in the event, $p_{T,i}$ the transverse momentum of the cluster and A_i the areas of the cluster.

A Constituent Subtraction (CS) algorithm which is a variant of the iterative pedestal subtraction technique [157, 158] is employed in this background subtraction. The CS subtraction is then performed jet-wise such that each jet constituent's 4-momentum is corrected. By correcting the jet constituents, both the full jet and jet substructure can be studied without the contamination of the background. Massless, low momentum

particles called "ghosts", are added to the event, uniformly in the $\eta - \phi$ plane such that each jet will contain both real and ghost particles with the total momentum of the ghosts in a jet given by,

$$\sum p_T^{g,k} = A_i \rho \qquad \ni \qquad p_T^{g,k} = A_g \rho \tag{3.10}$$

where A_g is the area of the ghost. A distance measure is defined for each pair of real particle i and ghost k:

$$\Delta R_{i,k} = p_{T,i} \sqrt{(\eta_i - \eta_k^g)^2 + (\phi_i - \phi_k^g)^2} \tag{3.11}$$

An iterative background removal procedure is performed for each pair of real and ghost particle modifying both their p_{T} and mass until all the particles are considered. The real or ghost particle with smaller p_{T} is subtracted from the larger and the smaller p_{T} particle is discarded at each step. Finally, the 4-momentum of the jet is recalculated using the corrected constituent particles [158]. Figure 3.23 shows the performance of the UE subtraction procedure in HYDJET MC relying on the estimate of the background density. The background fluctuations in the subtracted jet are then dependent on the radius of the jet. The accuracy of the background density is estimated by drawing random cones in the unsubtracted event where the p_{T} of all the particles is added and ρA is subtracted. It is seen that the most probable value of the distributions is close to zero indicating that the background density is correctly estimated. The residual background fluctuations are then corrected using unfolding for the substructure observables as discussed in Section 4.7.

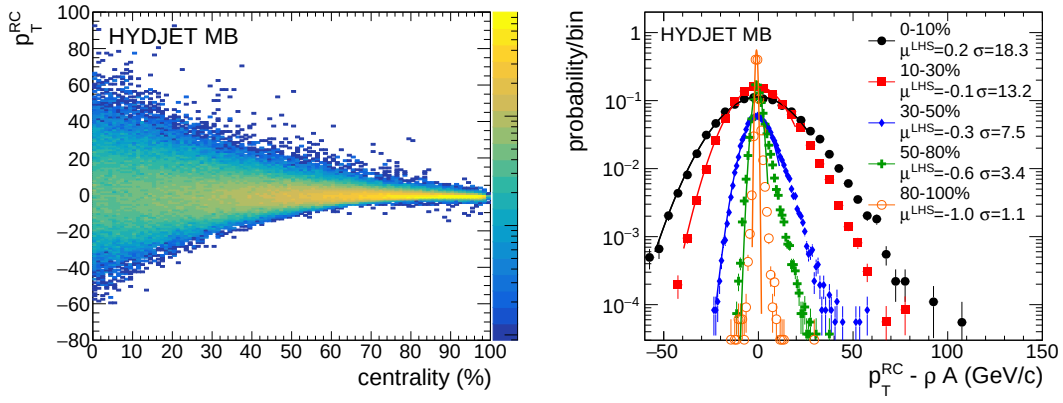


Figure 3.23: Performance of UE subtraction in HYDJET MC using random cones of radius $R = 0.4$. The distribution of the p_{T} of the random cone as a function of centrality (left) and and the most probably value(right) is shown in the figures adapted from Internal AN-16-069 [155]. The most probable value is close to zero indicating that the background density is correctly estimated.

3.3.2 Jet calibrations

High transverse momentum jets are generally associated with partons from the initial hard scattering. However, the reconstruction procedure does not necessarily recover the true energy and momentum, such that it can be used as a proxy for the parton in analyses. This necessitates obtaining jet energy corrections through a combination of simulated MC and data comparisons. The jets are corrected for the detector response with jet energy scale corrections (JES or JEC) derived from independent PbPb and pp simulations along with additional corrections for the imperfect modeling of the detector response [159]. The jet energy corrections are obtained for pp collisions [160] followed by corrections to account for PbPb collision environments. In addition to the JEC, the momentum resolution of jets is larger in data compared with simulation. Corrections to the simulation to account for the momentum resolution differences are derived from dijet balancing studies performed in pp collisions at $\sqrt{s} = 13\,\mathrm{TeV}$ in 2017 and 2018 [160].

The jet energy scale is the correction to the energy of the jet, obtained through a factorized approach that on average restores the energy of the reconstructed jet to the particle level energy [160]. The JECs are obtained in 3 stages, (i) "L1" which corrects

for the pileup, (ii) "L2" which are corrections obtained through simulated responses, and (iii) "L3" which are residual corrections obtained through comparisons between simulation and data. Due to the low pileup environments in the reference pp as well as PbPb collisions, pileup subtraction methods are not implemented with the focus being solely on the UE background subtraction in the case of PbPb as discussed in Section 3.3.1. Thus, the derivation of L2L3 corrections are discussed in the context of heavy ion collisions.

The study of simulated responses for JEC is possible through GEANT4 [161] which simulates the interaction of electromagnetic and hadronic showers using the detailed detector geometry of CMS. Together with parton fragmentation models in MC such as PYTHIA and HERWIG, an accurate jet response in the detector is obtained with small residual corrections based on data obtained as a further correction. Some advantages of using simulated responses in obtaining the JEC are the lack of biases inherent to using data-driven methods and a wider phase space coverage than is available easily in data. The jet energy response (R_{ptcl}) is defined as the ratio of arithmetic means of matched reconstructed and particle-level jets transverse momenta [160],

$$R_{\mathrm{ptcl}}(\langle p_{\mathrm{T}}\rangle, \eta) = \frac{\langle p_{\mathrm{T}}\rangle}{\langle p_{T,\mathrm{ptcl}}\rangle}[p_{T,\mathrm{ptcl}}, \eta] \tag{3.12}$$

in bins of particle level p_{T}, ($p_{T,\mathrm{ptcl}}$) and reconstructed η.

The data driven jet energy corrections are obtained to correct the relative and absolute JES using dijet, Z+jet and γ +jet events in data and simulation. The relative JES is determined using dijet p_{T} balancing method. The p_{T} balancing method requires one barrel jet to be in the central region ($|\eta| < 1.3$) with another probe jet at arbitrary η. The balance is then defined as,

$$\mathcal{B} = \frac{p_{\mathrm{T}}^{\mathrm{probe}} - p_{\mathrm{T}}^{\mathrm{barrel}}}{p_{\mathrm{T}}^{\mathrm{average}}} \tag{3.13}$$

where $p_{\mathrm{T}}^{\mathrm{average}}$ is the average p_{T} of the two jets. The average balance $\langle \mathcal{B} \rangle$ is then used to determine the relative response R_{rel} as,

$$R_{\mathrm{rel}} = \frac{2 + \langle \mathcal{B} \rangle}{2 - \langle \mathcal{B} \rangle} \equiv \frac{\langle p_{\mathrm{T}}^{\mathrm{probe}} \rangle}{\langle p_{\mathrm{T}}^{\mathrm{barrel}} \rangle} \tag{3.14}$$

where the equivalence holds for narrow bins of $p_{\mathrm{T}}^{\mathrm{average}}$.

The absolute JES is determined using Z+jet and γ +jet events with a complementary method known as the missing transverse momentum projection fraction (MPF) [159] and verified with the p_{T} balancing method. The MPF method is considered the response of the full hadronic activity in the event by using the fact that at parton level, the γ or Z is perfectly balanced by the hadronic recoil in the transverse plane and have no intrinsic missing transverse momentum ($E_{\mathrm{T}}^{\mathrm{miss}}$). This gives a response for the recoil jets (R_{recoil}),

$$R_{\mathrm{recoil}} = R_{\gamma,\mathrm{Z}} + \frac{E_{\mathrm{T}}^{\mathrm{miss}} p_{\mathrm{T}}^{\gamma,\mathrm{Z}}}{(p_{\mathrm{T}}^{\gamma,\mathrm{Z}})^2} \equiv R_{\mathrm{MPF}} \tag{3.15}$$

The p_{T} balance in the case of γ/Z +jet events is determined to determine the absolute response,

$$R_{\mathrm{abs}} = \frac{p_{\mathrm{T}}^{\mathrm{jet}}}{p_{\mathrm{T}}^{\gamma,\mathrm{Z}}} \tag{3.16}$$

The JES using the p_{T} balancing method in pp collisions is determined relative to a precise measure for muons, electrons and photons, considering a tracker scale uncertainty of 0.2% for muons at $|\eta| < 2.4$ [162], ECAL energy scale uncertainty of 0.5% for electrons at $|\eta| < 2.4$ [163] and 0.2% for photons at $|\eta| < 1.3$ [141]. In the case of PbPb collisions, a factorized multi-step approach is used to parameterize the JES [156, 164]. Corrections are derived in first obtained in PYTHIA simulations without the UE and then validated with a combined PYTHIA+HYDJET sample. The resulting JES depends

on the p_{T} and centrality of the events.

The jet transverse momentum resolution or jet energy resolution (JER) is estimated in both data and MC using the dijet asymmetry and γ +jet balance methods. In the case of dijets, the asymmetry variable $\mathcal{A}$ and the variance on the asymmetry $\sigma^2_{\mathcal{A}}$ is defined as,

$$\mathcal{A} = \frac{p_{\mathrm{T}}^{\mathrm{jet1}} - p_{\mathrm{T}}^{\mathrm{jet2}}}{p_{\mathrm{T}}^{\mathrm{jet1}} + p_{\mathrm{T}}^{\mathrm{jet2}}} \qquad \sigma^2_{\mathcal{A}} = \left|\frac{\partial \mathcal{A}}{\partial p_{\mathrm{T}}^{\mathrm{jet1}}}\right|^2 \sigma^2(p_{\mathrm{T}}^{\mathrm{jet1}}) + \left|\frac{\partial \mathcal{A}}{\partial p_{\mathrm{T}}^{\mathrm{jet2}}}\right|^2 \sigma^2(p_{\mathrm{T}}^{\mathrm{jet2}}) \tag{3.17}$$

For jets in the same η region, the averages of the jet p_{T} and variances are equivalent for the 2 jets allowing the fractional jet p_{T} to be calculated as,

$$\frac{\sigma(p_{\mathrm{T}})}{p_{\mathrm{T}}} = \sqrt{2}\sigma_{\mathcal{A}} \tag{3.18}$$

The dijet asymmetry method can be used to estimate the resolution using two methods: by considering the RMS to characterize the core of the distribution and a Gaussian or double sided crystal ball fitting of the full jet p_{T} resolution. The γ +jet method, on the other hand exploits the balance of the jet with the photon in the transverse plane to accurately measure the width of the $p_{\mathrm{T}}^{\mathrm{jet}}/p_{\mathrm{T}}^{\gamma}$ distribution to determine the jet resolution. These methods serve as complimentary cross-checks for each other.

The particle level JER in simulation is parameterized with the "CSN" fit for calorimeter resolutions where the C is a constant term dependent on the inherent calorimeter calibrations, S is a stochastic term that scales as $1/\sqrt{E}$ and N is a noise term [160]. In the case of pp collisions, the noise term N is dominated by electronic noise and pileup, while in PbPb collisions, it is dependent on the UE background. Thus, the treatment of JER excludes the noise term with an independent study on the effect of underlying event [165].

Chapter 4

Analysis procedure

This chapter provides the details about the procedure followed to perform the analysis motivated in Chapter 1. The analysis is performed by comparing the collisions in a heavy ion system (PbPb) consisting of lead ions to a proton system (pp). The data obtained from the CMS detector and MC samples used in this analysis are explained in Section 4.1. Section 4.2 discusses the principal difference between the PbPb and pp collisions, namely the centrality parameter. In Sections 4.3 and 4.4, further details about the selection criteria used for the photon and jet objects in this analysis are provided.

The observables of interest for this analysis, the groomed jet radius and the jet girth are defined in 4.5 followed by plots of photon and jet distributions used in the event selection and construction of these observables in both PbPb and pp collision systems in Section 4.6. These distributions are smeared by the detector and so they are corrected to their true distributions using an unfolding procedure described in Section 4.7. The systematic uncertainties considered in this analysis are described in Section 4.8.

4.1 Datasets

The analysis presented in this thesis uses the PbPb data collected in 2018 at (nucleon-nucleon) center of mass energy of 5.02 TeV with integrated luminosity of $1.7\,\mathrm{nb}^{-1}$ and the pp data at the same center of mass energy collected in 2017 with a luminosity of $301\,\mathrm{pb}^{-1}$ using the CMS detector.

The main processes of interest in this analysis are events composed of prompt photons and jets produced through the strong interaction. At leading order in perturbation theory, these are produced by $q\bar{q} \to \gamma g$ and $qg \to q\gamma$ as described in Section 3.2. Since the gluon parton distribution functions (PDF) are larger than the anti-quark PDFs (for the phase-space considered in the measurement), it is expected that $qg \to q\gamma$ dominates the sample in PbPb or pp collisions. The data collected by the CMS detector is stored in separate primary datasets (PD) depending on the output of a set of triggers decided upon during the data-taking process. The processes of interest for this analysis are in the `HiHardProbes` PD for PbPb and the `HighEGJet` PD for pp data.

To maximize the amount of data utilized in the analysis, the lowest threshold unprescaled trigger was used where events are required to pass the L1 trigger with E_T threshold of 21 GeV and to pass the HLT with a E_T threshold of 40 GeV. This corresponds to the following L1 seeds and HLT paths listed in Table 4.1. The triggers are fully efficient for offline photons $p_{\mathrm{T}}^{\gamma} > 100\,\mathrm{GeV}$ for both pp and PbPb collisions and require no corrections as shown in Section 5.4. Further CMS specific information on the datasets used in this analysis are provided in Appendix A for internal use.

Table 4.1: Level 1 (L1) and corresponding High-Level Trigger (HLT) paths used for pp and PbPb data

Collision system	L1	HLT
pp	`L1_SingleEG21`	`HLT_HIPhoton40_HoverELoose_v1`
PbPb	`L1_SingleEG21_BptxAND`	`HLT_HIGEDPhoton40_v1`

The study of these events in data requires simulated processes at various stages of the

analysis. Their simulation is performed with two main combinations of MC event generators to correct for the detector response and to extract systematic uncertainties. A sample of jet events is generated at leading-order (LO) with PYTHIA 8.230 [166], which implements a dipole shower ordered in p_T and where the hadronization of quarks and gluons into stable hadrons is described by the Lund string model [12, 167]. The PYTHIA8 parameters for the underlying event are set according to the CP5 tune [168]. For PbPb collisions, one has to consider the fact that there is not only one nucleon-nucleon collision (the one responsible for the hard scattering) but multiple nucleon-nucleon collisions occurring at the same time. Thus, multiple nucleon-nucleon scatterings simulated in PYQUEN based on PYTHIA6 and including quenching effects are superimposed on a hydrodynamical background simulated by a heavy ion generator like HYDJET v1.9, and then reconstructed like normal events. This step is needed because the detector response is affected by its occupancy degrading the detector performance for higher multiplicity events as is the case in central collisions. For pp collisions, a second sample is generated at LO with HERWIG 7.2.2 with the CH3 underlying event tune [169] to assess systematic uncertainties related to the modeling of the parton shower and hadronization. In HERWIG7, the parton shower follows angular-ordered radiation [170], and the hadronization is described by the cluster fragmentation model [13]. The LO NNPDF 3.1 [171] PDFs with $\alpha_S(m_Z) = 0.118$ are used, where m_Z is the Z boson mass. All generated samples are passed through a detailed simulation of the CMS detector using GEANT4 [161]. The MC simulated events are reweighted to reproduce the distribution of the number of interactions observed in the experimental data.

4.2 Centrality determination

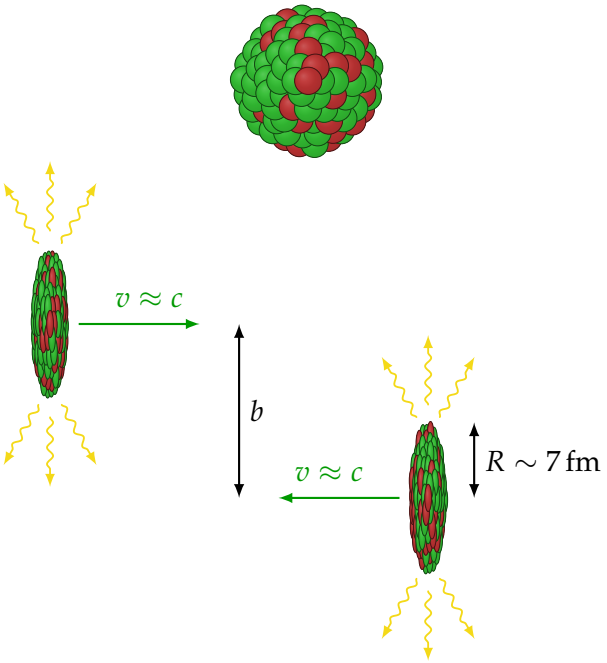


Figure 4.1: Schematic view of the initial geometry of PbPb collisions. At rest frame the nucleon is spherical as shown (upper). The lead ions are then flattened into an elliptical shape due to lorentz contraction while colliding at $v \approx c$ (lower). Adapted from Ref.[129]

The size and evolution of the matter created in a relativistic heavy-ion collision strongly depend on collision geometry, defined by centrality. In theory, the collision centrality C_b is defined as a fraction (expressed as a percentage) of the total inelastic nucleus–nucleus (AA) cross-section, $\sigma^{AA}_{inel.}$:

$$C_b = \frac{1}{\sigma^{AA}_{inel.}} \int_0^b \frac{d\sigma}{db'} db' \tag{4.1}$$

Here, b is the impact parameter and $\frac{d\sigma}{db'}$ is the differential cross-section of AA collision. The impact parameter b, defined as the transverse distance between the centers of the two colliding nuclei, is a well-defined quantity and a key input to most theoretical calculations of heavy-ion collisions as shown in Fig. 4.1. However, one cannot directly measure the impact parameter. Experimentally, the heavy-ion collisions are instead characterized by the measured multiplicity N_{ch} of the produced charged particles

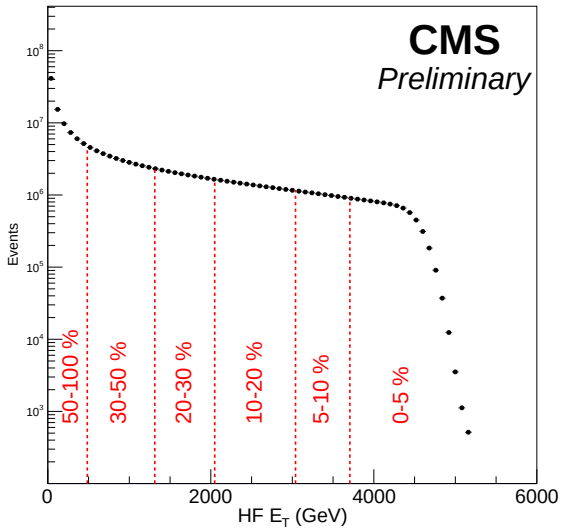


Figure 4.2: The sum of transverse energy deposited in the HF calorimeters is used for determining the centrality. A percentile distribution for minimum bias events is used to split the regions into centrality bins (shown in red) [165]

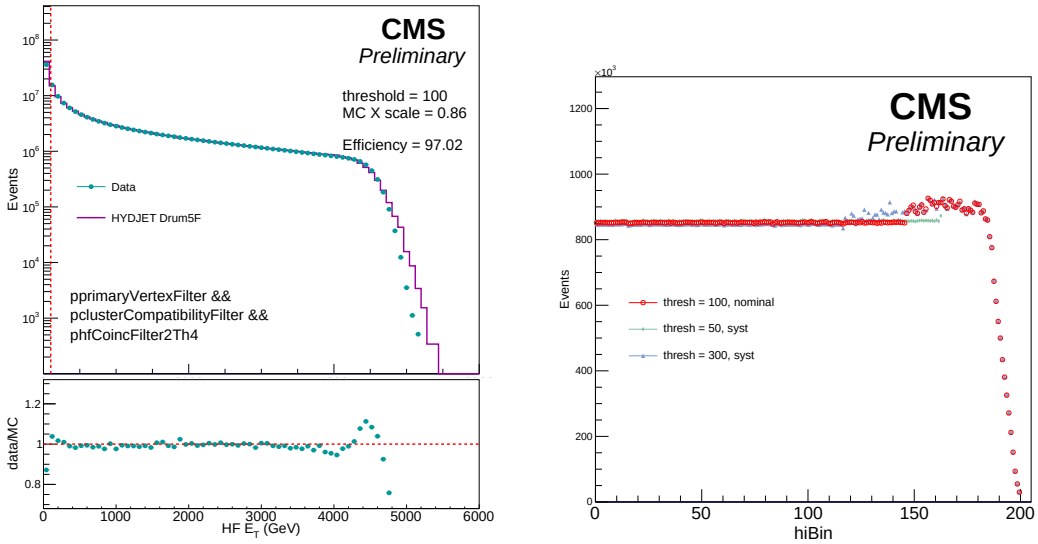


Figure 4.3: A comparison of the sum of transverse energy in the HF calorimeters between data and HYDJET MC is shown for a specified threshold(left). The HF energy is mapped to centrality classes of 0.5% corresponding to 200 bins are shown for minimum bias events in 2018 PbPb data [165]

around midrapidity or by the energy measured in the forward rapidity region.

Simulations can be used to correlate centrality, as quantified using the fraction of the total interaction cross section, with more detailed properties of the collision. Two commonly used physical quantities are the total number of nucleons in the lead (^{208}Pb) nuclei which experienced at least one inelastic collision, denoted N_{part}, and the total number of binary inelastic nucleon-nucleon collisions N_{coll}. The centrality bins can be correlated to the impact parameter b, and to average values and variances of N_{part} and N_{coll} using a calculation based on the Glauber model in which the nucleons are assumed to follow straight-line trajectories as they collide as described in section 1.3.3. Equation 4.2 follows from equation 1.20 using the Wood-Saxon distribution for the nuclear overlap function discussed to define N_{coll}. The bin-to-bin smearing of the results of these calculations due to the finite resolution and fluctuations in the hadron forward (HF) calorimeter transverse energy measurement was obtained from fully simulated and reconstructed MC events.

$$N_{coll}(b) = ABT_{AB}(b)\sigma_{nn}^{\text{inel}} \tag{4.2}$$

The centrality in PbPb collisions is determined from the total transverse energy deposited in the hadron forward (HF) calorimeters covering the forward region $2.9 < |\eta| < 5.2$ at both sides of the interaction point. The collision centrality is expressed as a percentage of the total inelastic hadronic cross-section, with 0% representing the most head-on (central) collisions and 100% the most peripheral collisions [82] as shown in Figure 4.2. Centrality classes of 0.5% of the total inelastic hadronic cross-section (obtained from minimum bias collisions) corresponding to 200 centrality bins are created for defined HF E_T boundaries. Various thresholds are compared for the HF energy in data as shown in Figure 4.3. Hadronic events can be selected by requiring hits in the forward calorimeter towers with specified energy thresholds.

4.3 Event weighting and quality cuts

The MC samples produced following the default settings do not necessarily describe the data accurately for the analysis needs. Event wise weights are applied to match the global quantities in data expected to be well described by the MC. Some checks are performed to verify this as described in Section 5.1. The MC samples used in this analysis are produced to enhance the presence of prompt photons. The cross-section of these processes depend strongly on the p_{T} of the parton and is a steeply falling spectrum. Thus, the MC samples are produced in multiple biased bins of $\hat{p_{\mathrm{T}}}$ to enhance higher p_{T} event yields for the analysis. This is corrected by applying a weight factor while combining the different MC samples as shown in Figure 4.4. In the case of PbPb MC samples, an additional weight to correct for the centrality is also applied.

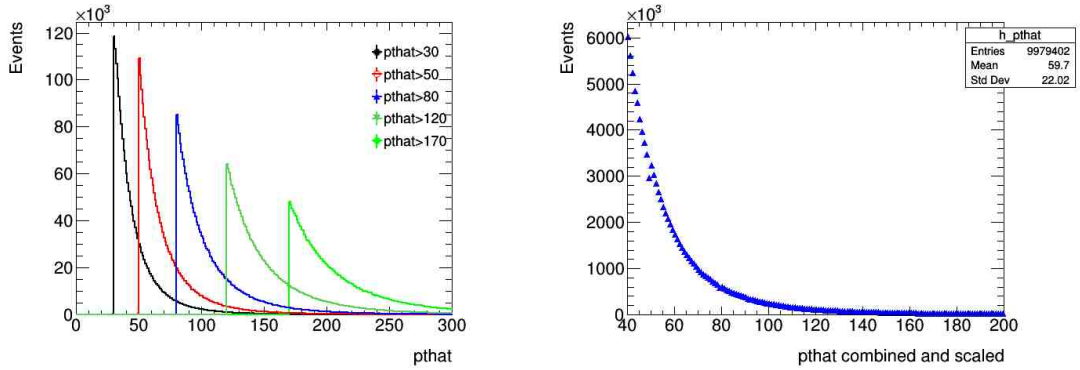


Figure 4.4: $\hat{p_{\mathrm{T}}}$ distributions (left) for the different MC samples used in this analysis and the reweighted $\hat{p_{\mathrm{T}}}$ distribution (right) after combining the samples

A standard offline event selection is applied to select hadronic events (by requiring a coincidence of two towers on each side of the interaction point of the HF calorimeters, with a threshold above 4 GeV), and to remove overall non-collisions events (i.e., beam-scraping events, beam-gas interactions, cosmic ray showers, etc.) by requiring a cluster compatibility and a valid primary vertex in both PbPb and pp collision data. Further, ECAL and HCAL noise filters are also applied.

4.4 Event Selection

This analysis is restricted to the 30% most central events in order to focus on medium modification effects in the case of PbPb events. Also, in PbPb events an additional veto region is applied due to the failure of the HCAL in $\eta < -1.39$ and $-1.6 < \phi < -0.9$ for both photons and jets. The analysis requires the presence of an isolated photon with $p_{\mathrm{T}}^{\gamma} > 100\,\mathrm{GeV}$ and $|\eta_{\gamma}| < 1.44$ and the presence of at least one anti-k_{T} jet with $R = 0.2$ in both PbPb and pp collision systems as shown in Figure 4.5. The jet selection is controlled using the transverse momentum imbalance, defined as the ratio of the p_{T} of the highest-p_{T} jet and the photon momentum $x_{\gamma \mathrm{j}} \equiv p_{\mathrm{T}}^{\mathrm{jet}} / p_{\mathrm{T}}^{\gamma}$. We considered two jet selections, such that the jet carries at least 40% of the photon momentum corresponding to $x_{\gamma \mathrm{j}} > 0.4$, as well as a stricter selection requiring the jet to carry at least 80% of the photon momentum. Both selections require $|\eta_{\mathrm{jet}}| < 2$ and an angular separation in azimuth defined by $\Delta\varphi_{\gamma j} > \frac{2}{3}\pi$. The $x_{\gamma \mathrm{j}} > 0.4$ requirement is such that jets have a minimum p_{T} of 40 GeV, while the $x_{\gamma \mathrm{j}} > 0.8$ selection is to account for the average p_{T} shift induced by out-of-cone radiation expected for jets produced in pp collisions [172]. The anti-k_{T} distance parameter $R = 0.2$ and threshold $p_{\mathrm{T}}^{\mathrm{jet}} > 40\,\mathrm{GeV}$ (from the $x_{\gamma \mathrm{j}} > 0.4$ selection) are chosen such that the reconstructed jet at detector-level bears a strong correspondence with the corresponding hard jet at particle-level, and reduces the likelihood that the jet originates from an UE fluctuation to allow for robust unfolding performance. The substructure observables are measured using the highest p_{T} jet recoiling from the selected isolated photon. The large UE in PbPb collisions can create spurious structures at the level of the jet substructure. This is mitigated with the use of small $R = 0.2$ jets, since the UE tends to manifest for large splitting angles. In addition, soft-drop grooming with $z_{\mathrm{cut}} = 0.2$ is used to further suppress these contributions. The reason is that splittings induced by the large UE tend to have a more asymmetric momentum balance than hard splittings from a genuine hard parton branching process at large angles. For $p_{\mathrm{T}}^{\gamma} > 100\,\mathrm{GeV}$, triggers are fully efficient both in pp and PbPb systems. No addi-

tional correction to account for residual trigger inefficiencies is applied. The number of selected γ +jet pairs with $x_{\gamma j} > 0.4$ ($x_{\gamma j} > 0.8$) is 4717 (1940) in PbPb collisions. In pp collisions, the number of reconstructed events with $x_{\gamma j} > 0.4$ ($x_{\gamma j} > 0.8$) is 20636 (10796).

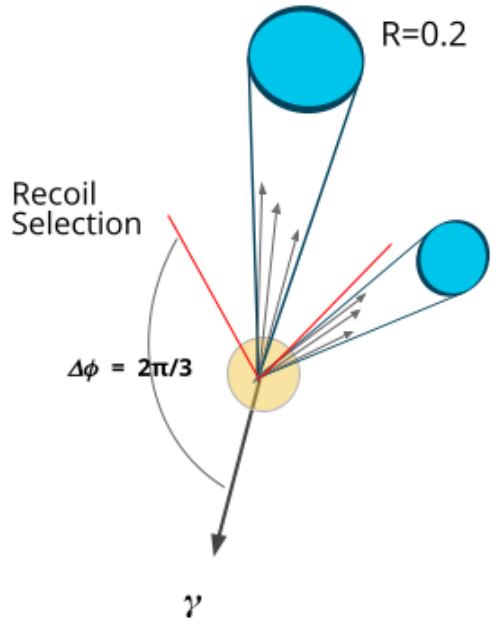


Figure 4.5: Schematic diagram showing the event selection with a photon and the leading recoiling jet

For parton energy loss studies in heavy-ion collisions, a selection of direct photons is desired since they provide a reliable proxy for the momentum of the recoil parton that initiated the jet shower. To suppress the contribution from the background of decay and fragmentation photons, a cut-based approach using a multivariate analysis [143] is performed as described in Section 3.2. The signal at particle-level are photons that are isolated. At particle level, the isolation criterion is such that the scalar sum of the p_{T} of all the particles around the photon within a cone of $\Delta R < 0.4$ should be less than 5 GeV. The detector-level variables used in the multivariate analysis are the fraction of hadronic energy around the photon candidate, the shower shape variable $\sigma_{\eta\eta}$ and the isolation variable $\mathcal{I}$ [122]. A working point corresponding to a signal efficiency of 60%

is chosen in both PbPb and pp collisions as obtained from simulation in Section 3.2 as follows:

PbPb:

- H/E $<$ 0.119947
- $\sigma_{\eta\eta}$ $<$0.010392
- $\mathcal{I} < 2.099277$

pp:

- H/E $<$ 0.009732
- $\sigma_{\eta\eta} < 0.009905$
- $\mathcal{I} < -0.014755$

These cuts provide a relatively pure sample of isolated photons. A residual contribution originating from misidentified neutral meson decays are a dominant source of background for prompt photons. The residual contribution of these decays is taken into account via signal photon purity of the cut-based selected sample. A binned maximum likelihood fit is performed using a signal template obtained from MC and a background template from nonisolated photons for the nominal correction. An alternative data-driven ABCD method is also performed as discussed in Section 3.2. The estimated photon purities are consistently obtained with the two methods and have values of 0.93 ± 0.02 and 0.77 ± 0.01 for pp and PbPb events, where the central value corresponds to the photon purity obtained with the template fit method and the uncertainties correspond to the symmetrized difference of the photon purity obtained with the alternative ABCD method described in Section 4.7. According to simulations, the selected sample of signal photons used in this analysis is dominated by direct photons with a background of approximately 15% of fragmentation photons that contribute to the tails of the $x_{\gamma \mathrm{j}}$ distribution.

4.5 Observables

Transverse momentum imbalance is defined as the ratio of the p_{T} of the highest-p_{T} jet and the photon momentum $x_{\gamma \mathrm{j}} \equiv p_{\mathrm{T}}^{\mathrm{jet}} / p_{\mathrm{T}}^{\gamma}$. In the case of inclusive jet measurements, an alternative momentum imbalance known as dijet asymmetry, $A_J = \frac{p_{\mathrm{T}}^1 - p_{\mathrm{T}}^2}{p_{\mathrm{T}}^1 + p_{\mathrm{T}}^2}$ is used, where the subscript 1 always refers to the leading jet so A_J is positive by construction with the subscript 2 defining the subleading jet [82]. The momentum imbalance $x_{\gamma \mathrm{j}}$ is used as a proxy for jet energy loss since photons do not interact strongly with the QGP medium.

The groomed jet radius R_{g} is the angle between the two subjets selected by the soft drop grooming algorithm. The grooming algorithm consists of first reclustering the jet constituents obtained from the anti-k_{T} jet [18] using the Cambridge–Aachen algorithm to impose angular ordering [19]. Secondly, the algorithm iterates back through the new clustering and selects the first pair of subjets that satisfy the condition $z > z_{\mathrm{cut}} \theta^{\beta}$. Here, z corresponds to the momentum fraction, $z = p_{\mathrm{T}}^{\mathrm{sub}} / (p_{\mathrm{T}}^{\mathrm{sub}} + p_{\mathrm{T}}^{\mathrm{lead}})$ where $p_{\mathrm{T}}^{\mathrm{lead}}$ ($p_{\mathrm{T}}^{\mathrm{sub}}$) is the momentum of the harder (softer) subjet in the selected pair and θ the splitting angle in rapidity-azimuth between the harder and softer subjets. In this analysis, we use the parameters $z_{\mathrm{cut}} = 0.2$ and $\beta = 0$ for soft-drop grooming, which allows us to better control the UE background in PbPb collisions [173]. The $z_{\mathrm{cut}} = 0.2$ and $\beta = 0$ selection was also used by the ALICE and ATLAS Collaborations [81, 96], so this choice also enables a more direct comparison across experiments.

In addition to R_{g}, we measure another infrared- and collinear-safe observable, the jet girth g [79, 174–176], which is defined as the sum of the products of the momentum fraction of the jet constituents and their distance relative to the anti-k_{T} jet axis, namely:

$$g = \frac{1}{p_{\mathrm{T}}^{\mathrm{jet}}} \sum_i p_{\mathrm{T}}^i \Delta R_{i,\mathrm{jet}}, \tag{4.3}$$

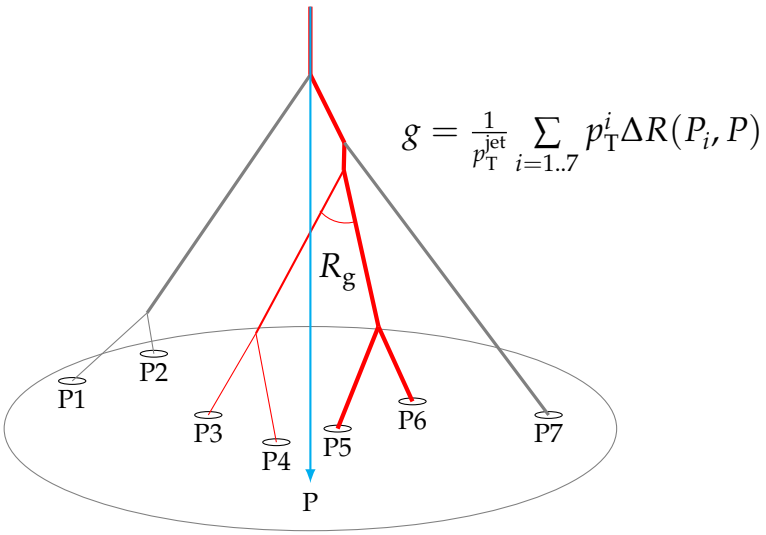


Figure 4.6: Schematic showing the observables groomed radius (R_{g}) and girth (g) of jets. The jet constituents are labelled P_i with the grooming algorithm iteration steps shown in red, thicker line. The jet axis is the blue, thin line with an arrow labelled P. The R_{g} is the angle between the subjets that satisfy the soft-drop condition, while the girth is the sum of the products of the momentum fraction of all the jet constituents and their distance to the jet axis

where $\Delta R_{i,\mathrm{jet}} = \sqrt{(\Delta y_{i,\mathrm{jet}})^2 + (\Delta \phi_{i,\mathrm{jet}})^2}$ is the distance in rapidity-azimuth of the i-th jet constituent with respect to the anti-k_{T} jet axis. Figure 4.6 is a schematic representation of the observables. The observables g and R_{g} are derived in different ways starting from the constituents of the anti-k_{T} jet. Since g does not rely on organizing radiation into clusters of particles as R_{g} does, and since it uses all hadrons in the jet, it potentially highlights different aspects of the jet fragmentation in vacuum and in the medium. Applying selections on $x_{\gamma \mathrm{j}}$ in conjunction with these observables allows for a controlled measurement on the medium modification effects of jets.

4.6 Raw detector-level distributions

The CMS detector can be sensitive to certain particles such as photons, electrons or jets during the data taking period causing hotspots to appear in regions of the η-ϕ plane. To verify that our analysis selection excludes such regions, we prepared 2D maps for the photon and jets. Fig. 4.7 shows the distributions for 0–30% central PbPb collisions. Fig. 4.8 shows the distributions for pp collisions. Figs. 4.9, 4.10 and 4.11 show the jet p_{T}, η and ϕ distributions in data and MC showing that they are well described by the models. Since quenching effects are not included in the MC, a shift is seen in the jet p_{T} distribution for central PbPb when compared to PYTHIA embedded MC in Fig. 4.9.

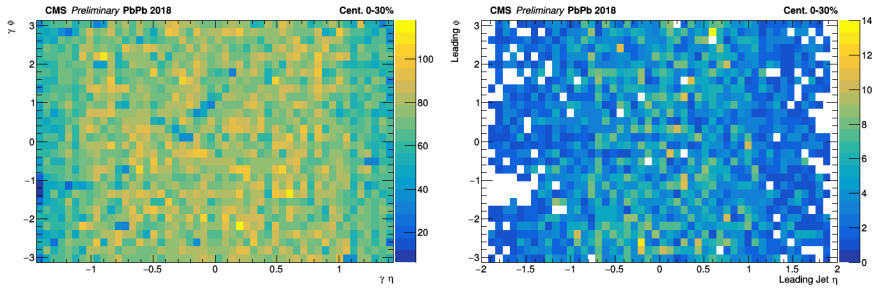


Figure 4.7: Photon (left) and jet (right) distribution in η-ϕ for 0–30% PbPb collisions for the analysis cuts. The plots show a uniform distribution as expected without any hotspots with the expected depletion due to the veto region accounting for the HCAL detector failure.

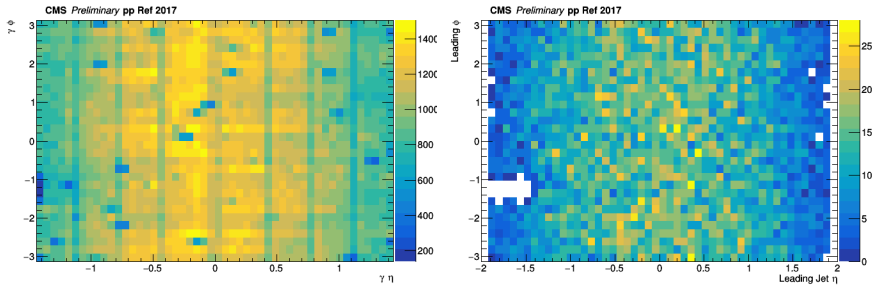


Figure 4.8: Photon (left) and jet (right) distribution in η-ϕ for pp collisions for the analysis cuts. The plots show a uniform distribution as expected without any hotspots.

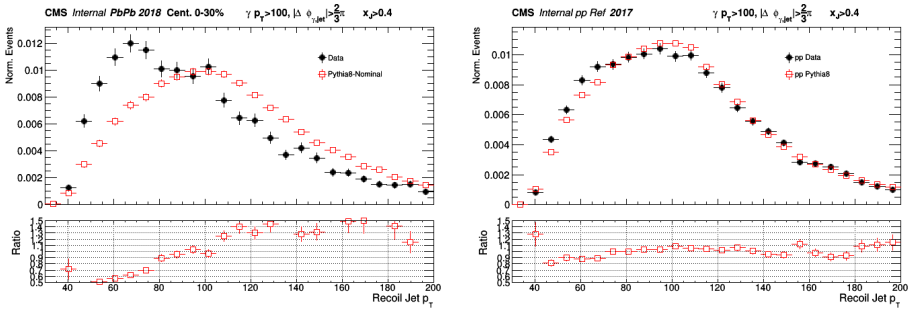


Figure 4.9: Jet p_T Distributions for 0–30% PbPb (left), and pp Data (right) compared to PYTHIA8 (Embedded) MC for the analysis cuts.

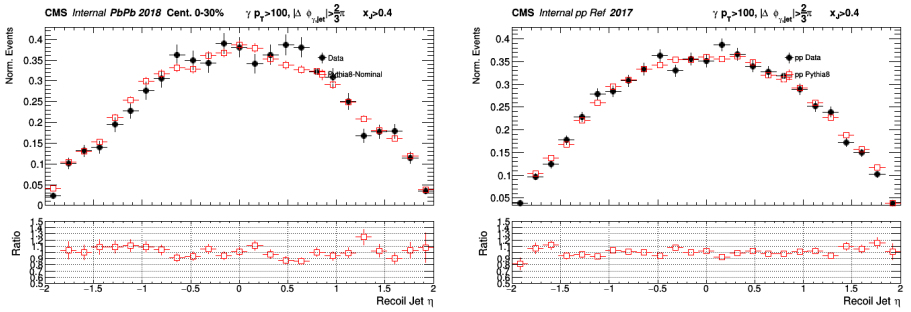


Figure 4.10: Jet η Distributions for 0–30% PbPb (left), and pp Data (right) compared to PYTHIA8 (Embedded) MC for the analysis cuts.

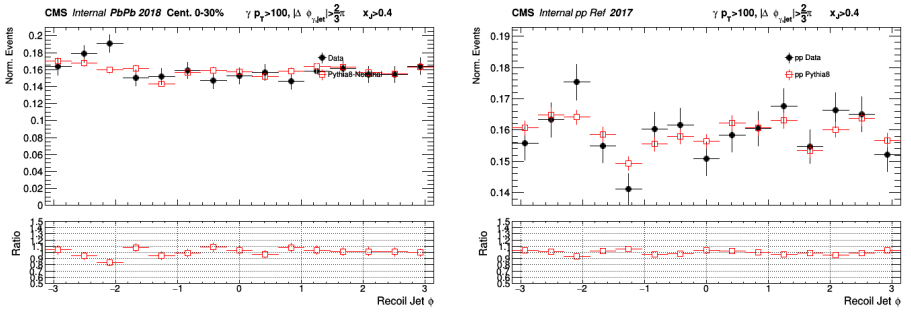


Figure 4.11: Jet ϕ Distributions for 0–30% PbPb (left), and pp Data (right) compared to PYTHIA8 (Embedded) MC for the analysis cuts.

4.6.1 Comparisons between MC and pp data at detector level

The raw-level distributions are compared to MC at detector level in Figs. 4.12, 4.13 and 4.14 in pp collisions for both $x_{\gamma j} > 0.4$ and 0.8 for $x_{\gamma j}$, girth g and R_g, respectively. The HERWIG7 describes the substructure distributions in data better than PYTHIA8. The distributions are normalized to the jets satisfying the respective $x_{\gamma j}$ condition.

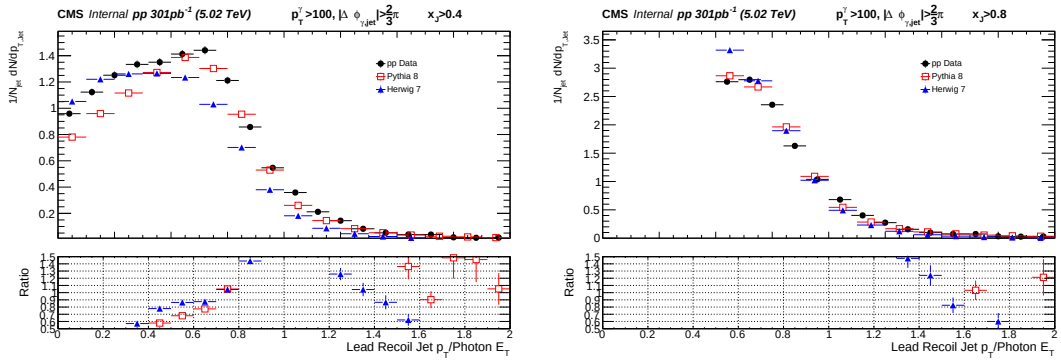


Figure 4.12: Detector-level transverse momentum imbalance $x_{\gamma j} \equiv p_{\mathrm{T}}^{\mathrm{jet}} / p_{\mathrm{T}}^{\gamma}$ distribution compared between pp data to the PYTHIA8 and HERWIG7 MC for $x_{\gamma j} > 0.4$ (left) and $x_{\gamma j} > 0.8$ (right).

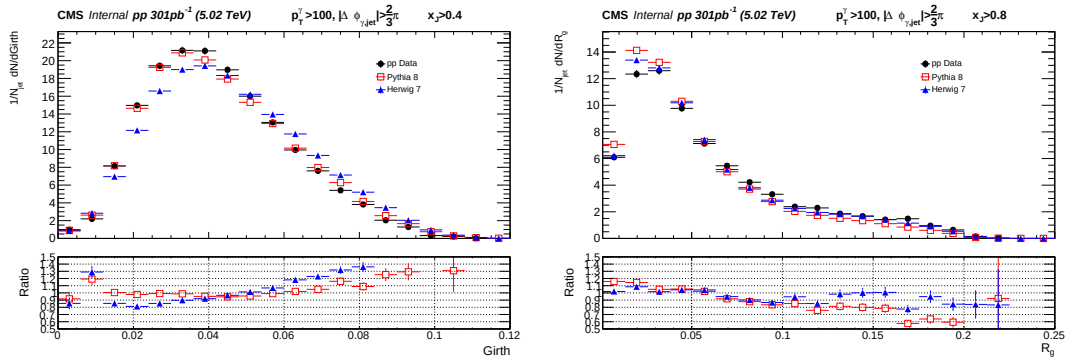


Figure 4.13: Detector-level jet girth g distribution compared between pp data to the PYTHIA8 and HERWIG7 MC for $x_{\gamma j} > 0.4$ (left) and $x_{\gamma j} > 0.8$ (right).

4.6.2 Comparisons between MC and PbPb data at detector level

The raw-level distributions are compared to MC at detector level in Figs. 4.15, 4.16 and 4.17 in central PbPb collisions for $x_{\gamma j} > 0.4$ and 0.8 for $x_{\gamma j}$, girth g and R_g, respectively.

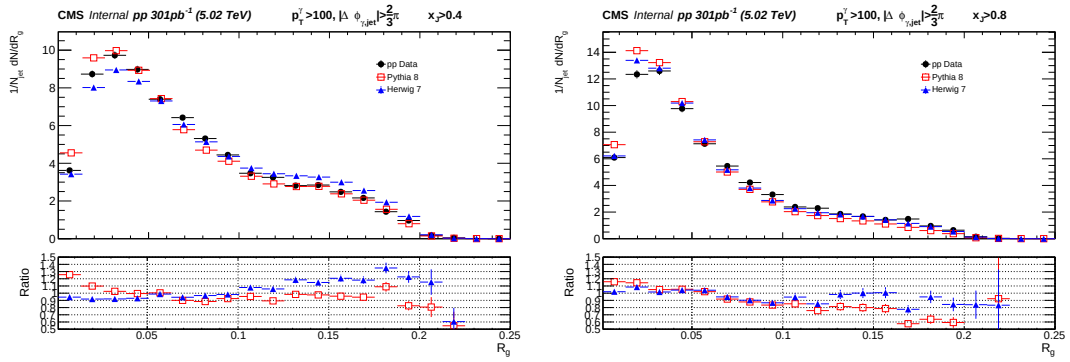


Figure 4.14: Detector-level groomed jet radius R_g (right) distribution compared between pp data to the PYTHIA8 and HERWIG7 MC for $x_{\gamma j} > 0.4$ (left) and $x_{\gamma j} > 0.8$ (right).

Significant differences are observed for the substructure distributions between PYTHIA8 embedded MC and the data. A template fitting method is used to reweight the quark-gluon fraction in the nominal PYTHIA8 MC to obtain a better description of the data as described in Section 5.7. This alternate MC is used to estimate the MC modeling uncertainty and is described in Section 4.7.

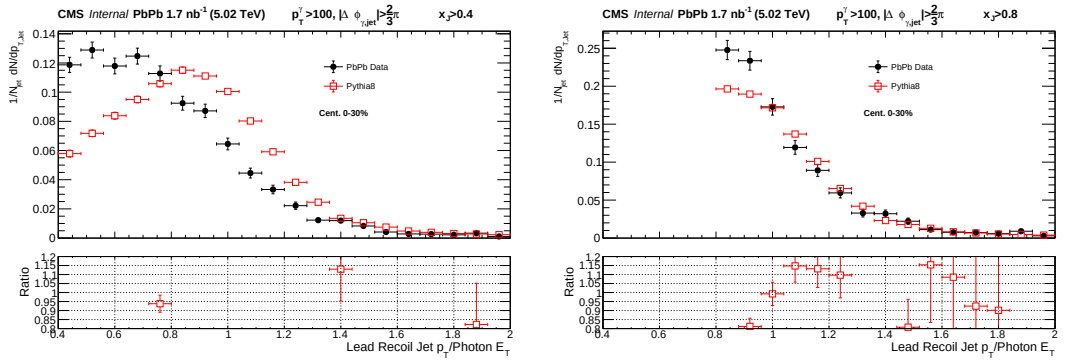


Figure 4.15: Detector-level transverse momentum imbalance $x_{\gamma j} \equiv p_{\mathrm{T}}^{\mathrm{jet}} / p_{\mathrm{T}}^{\gamma}$ distribution compared between 0–30% central PbPb data to the PYTHIA8 embedded MC for $x_{\gamma j} > 0.4$ (left) and $x_{\gamma j} > 0.8$ (right).

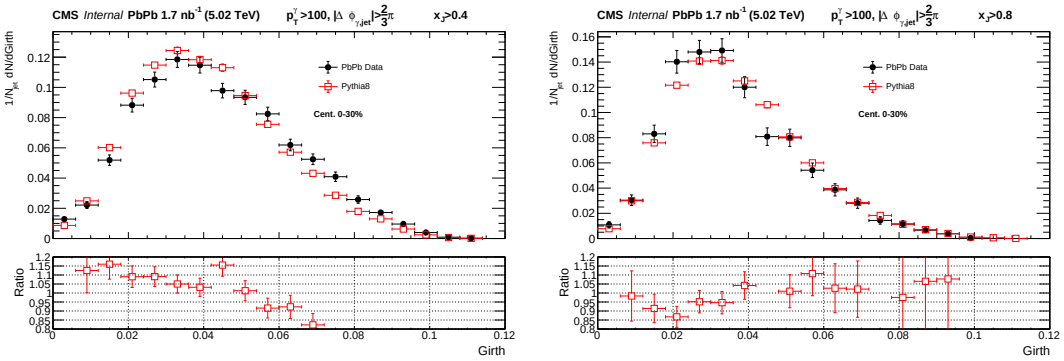


Figure 4.16: Detector-level jet girth g distribution compared between 0–30% central PbPb data to the PYTHIA8 embedded MC for $x_{\gamma j} > 0.4$ (left) and $x_{\gamma j} > 0.8$ (right).

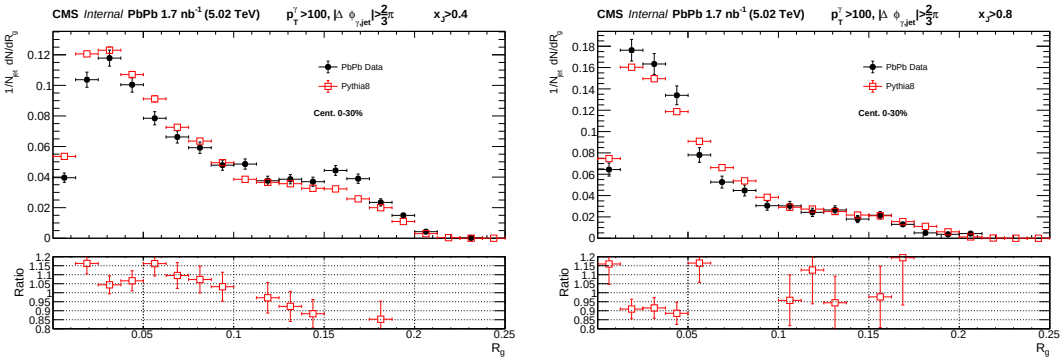


Figure 4.17: Detector-level groomed jet radius R_g distribution compared between 0–30% central PbPb data to the PYTHIA8 embedded MC for $x_{\gamma j} > 0.4$ (left) and $x_{\gamma j} > 0.8$ (right).

4.6.3 Comparisons between pp and PbPb at detector level

The photon p_{T} distribution is similar in both PbPb and pp collisions. This is expected as the photons are not expected to be modified by the medium and Fig. 4.18 illustrates this. The $x_{\gamma \mathrm{j}}$ comparison is shown in Fig.4.19. The strong modification of the variable is a proof of the strong quenching the recoiling jets undergo in PbPb collisions. We compare the jet girth g and groomed jet radius R_{g} distributions of the recoil jet in pp and PbPb at detector level. Comparisons of pp collisions to central PbPb collisions of the substructure observables are shown in Fig. 4.20 for g. Figure 4.21 shows the shape of the R_{g} distribution normalized to jets that satisfy the grooming condition. A count of the jets which failed the soft drop grooming condition is also kept. These are called untagged jets and treated separately when correcting the results to particle level as described in Section 4.7.

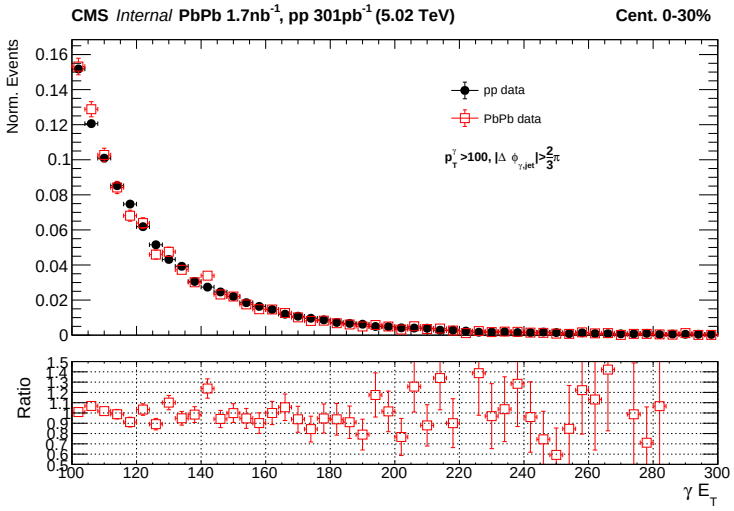


Figure 4.18: Detector-level photon transverse momentum p_{T}^{γ} distribution compared between 0–30% central PbPb data to pp data.

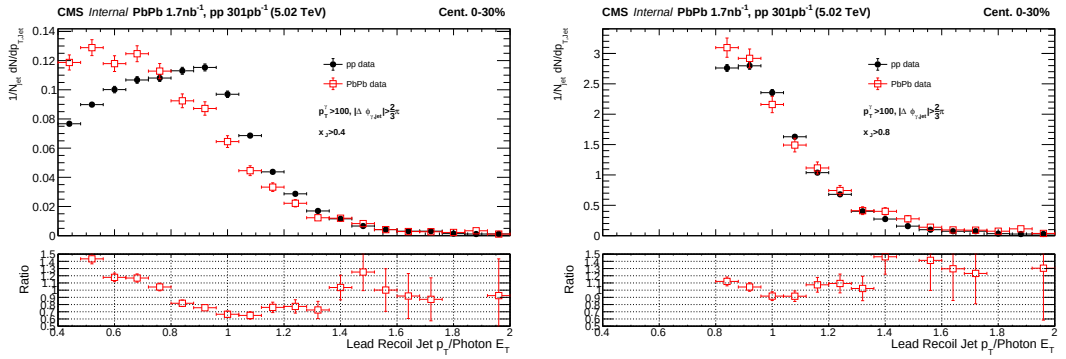


Figure 4.19: Detector-level transverse momentum imbalance $x_{\gamma j} \equiv p_{\mathrm{T}}^{\mathrm{jet}} / p_{\mathrm{T}}^{\gamma}$ distribution compared between 0–30% central PbPb data to pp data for $x_{\gamma j} > 0.4$ (left) and $x_{\gamma j} > 0.8$ (right). One can observe the effects of jet quenching in central PbPb collisions, as expected.

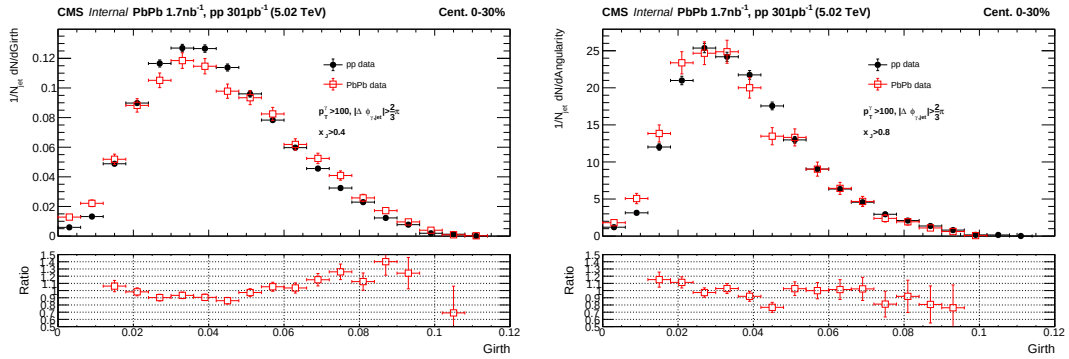


Figure 4.20: Detector-level jet girth g distribution compared between 0–30% central PbPb data to pp data for $x_{\gamma j} > 0.4$ (left) and $x_{\gamma j} > 0.8$ (right).

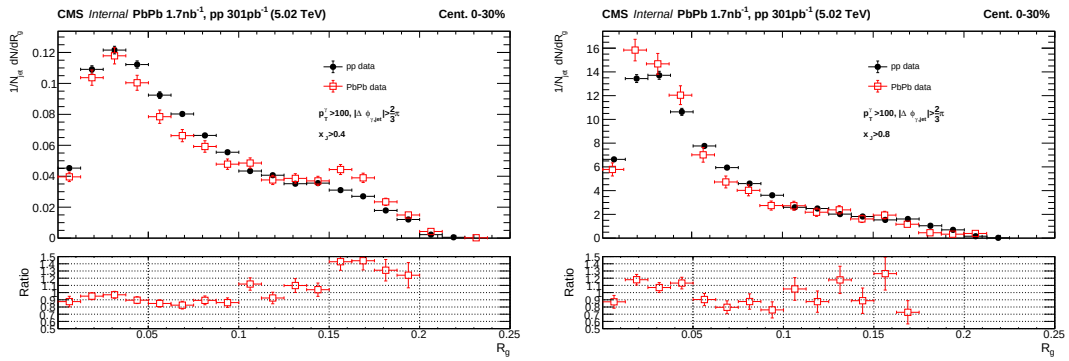


Figure 4.21: Detector-level groomed jet radius R_{g} distribution compared between 0–30% central PbPb data to pp data for $x_{\gamma j} > 0.4$ (left) and $x_{\gamma j} > 0.8$ (right).

4.7 Unfolding

In order to facilitate comparisons with other experiments and with theoretical predictions, we unfold the detector-level distributions to the particle level. The corrections account for efficiency, acceptance, and smearing effects. To quantitatively describe the migration effects, a multidimensional response matrix is calculated using isolated photon+jet events at particle and detector level. Both isolated photons and recoiling jets at particle and detector level are matched by proximity in η–ϕ space. The migrations in $x_{\gamma j}$ are dominated by the smearing of the jet p_{T} but also encode the migrations due to detector effects in the photon energy reconstruction.

Table 4.2: Bin-by-bin purity for Girth (g) for 0–30% PbPb and pp with $x_{\gamma j} > 0.4$ and $x_{\gamma j} > 0.8$ obtained using MC

Binning	**0–0.02**	**0.02–0.04**	**0.04–0.06**	**0.06–0.08**	**0.08–0.1**
PbPb $x_{\gamma j} > 0.4$	0.99	0.99	0.97	0.93	0.89
PbPb $x_{\gamma j} > 0.8$	1.00	1.00	1.00	0.99	0.97
pp $x_{\gamma j} > 0.4$	0.99	0.99	0.98	0.97	0.96
pp $x_{\gamma j} > 0.8$	1.00	1.00	1.00	1.00	0.98

Table 4.3: Bin-by-bin purity for R_{g} for 0–30% PbPb and pp with $x_{\gamma j} > 0.4$ and $x_{\gamma j} > 0.8$ obtained using MC

Binning	**untagged**	**0.00–0.04**	**0.04–0.08**	**0.08–0.12**	**0.12–0.20**
PbPb $x_{\gamma j} > 0.4$	0.98	0.99	0.98	0.96	0.93
PbPb $x_{\gamma j} > 0.8$	1.00	1.00	1.00	1.00	0.99
pp $x_{\gamma j} > 0.4$	0.99	0.99	0.98	0.97	0.97
pp $x_{\gamma j} > 0.8$	1.00	1.00	1.00	1.00	0.99

The feasibility of unregularized unfolding will be discussed in Section 5.10. Due to the instability of the unregularized unfolding results, regularized unfolding was determined to be required. Thus, we use the D'Agostini iterative unfolding with early stopping [177] as implemented in the RooUnfold package [178]. The unfolding is performed via a two-dimensional correction that simultaneously corrects the jet p_{T} via $x_{\gamma j}$ and the kinematics of the splittings (g or R_g). Bin-by-bin matching purity and efficiency corrections are applied before and after unfolding, respectively. The efficiency

and purity corrections are obtained from MC. The matching purity corrections account for the fraction of detector-level jets that are not assigned to a truth-level jet. The purity corrections are up to 10% in PbPb and up to 6% in pp, with the largest correction in the tail for highly quenched jets ($x_{\gamma j} > 0.4$) as shown in Tables 4.2 and 4.3. The efficiency corrections account for the fraction of particle-level jets that are not associated to a detector-level jet. The efficiency corrections are up to 20% in PbPb and pp, with the largest correction in the tail for less quenched jets ($x_{\gamma j} > 0.8$) as shown in Tables 4.4 and 4.5. The binning choice ensures a sufficiently large number of counts per-bin for stable unfolding corrections. The number of iterations, which plays the role of the regularization parameter in D'Agostini unfolding, is such that the unfolded solution folded back to detector level is statistically compatible with the input measured distribution.

Table 4.4: Bin-by-bin efficiency for Girth (g) for 0–30% PbPb and pp with $x_{\gamma j} > 0.4$ and $x_{\gamma j} > 0.8$ obtained using MC

Binning	**0–0.02**	**0.02–0.04**	**0.04–0.06**	**0.06–0.08**	**0.08–0.1**
PbPb $x_{\gamma j} > 0.4$	0.97	0.97	0.93	0.90	0.84
PbPb $x_{\gamma j} > 0.8$	0.91	0.87	0.84	0.79	0.74
pp $x_{\gamma j} > 0.4$	0.98	0.97	0.95	0.92	0.87
pp $x_{\gamma j} > 0.8$	0.90	0.87	0.85	0.79	0.74

Table 4.5: Bin-by-bin efficiency for R_g for 0–30% PbPb and pp with $x_{\gamma j} > 0.4$ and $x_{\gamma j} > 0.8$ obtained using MC

Binning	**untagged**	**0.00–0.04**	**0.04–0.08**	**0.08–0.12**	**0.12–0.20**
PbPb $x_{\gamma j} > 0.4$	0.95	0.97	0.94	0.91	0.86
PbPb $x_{\gamma j} > 0.8$	0.89	0.88	0.83	0.83	0.76
pp $x_{\gamma j} > 0.4$	0.98	0.97	0.95	0.92	0.89
pp $x_{\gamma j} > 0.8$	0.89	0.88	0.85	0.82	0.77

The stability of the unfolding is tested using the methods described in Section 5.11. The probability matrices for PbPb collisions are shown in Fig.4.22 and for pp collisions are shown in Fig. 4.23. The numbers in the bins represent the normalized probability for each truth bin. The binning also includes the underflow/overflow bins to account for out-of-phase migrations. The bin indices correspond to the bins used in the corresponding substructure distribution. In the PbPb case we observe some degree of non-

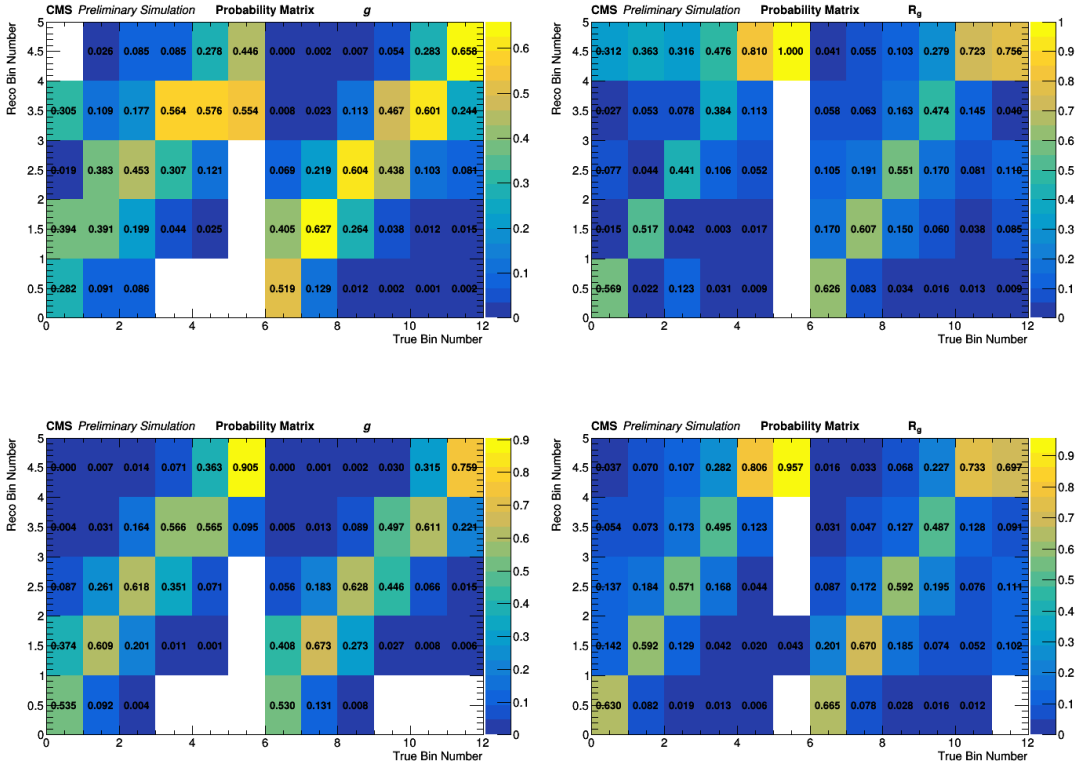


Figure 4.22: 2D Probability matrices for Girth (g) defined as (g True, g Reco) (left) and for R_g (right) defined as (R_g True, R_g Reco) for $x_{\gamma \mathrm{j}} > 0.4$ (top) and $x_{\gamma \mathrm{j}} > 0.4$ (bottom) in central PbPb simulation in PYTHIA8 embedded. The numbers in the bins represent the normalized probability for each truth bin. The binning also includes the underflow/overflow bins to account for out-of-phase-space migrations. The first 6 bins in the x-axis correspond to the underflow in $x_{\gamma \mathrm{j}}$.

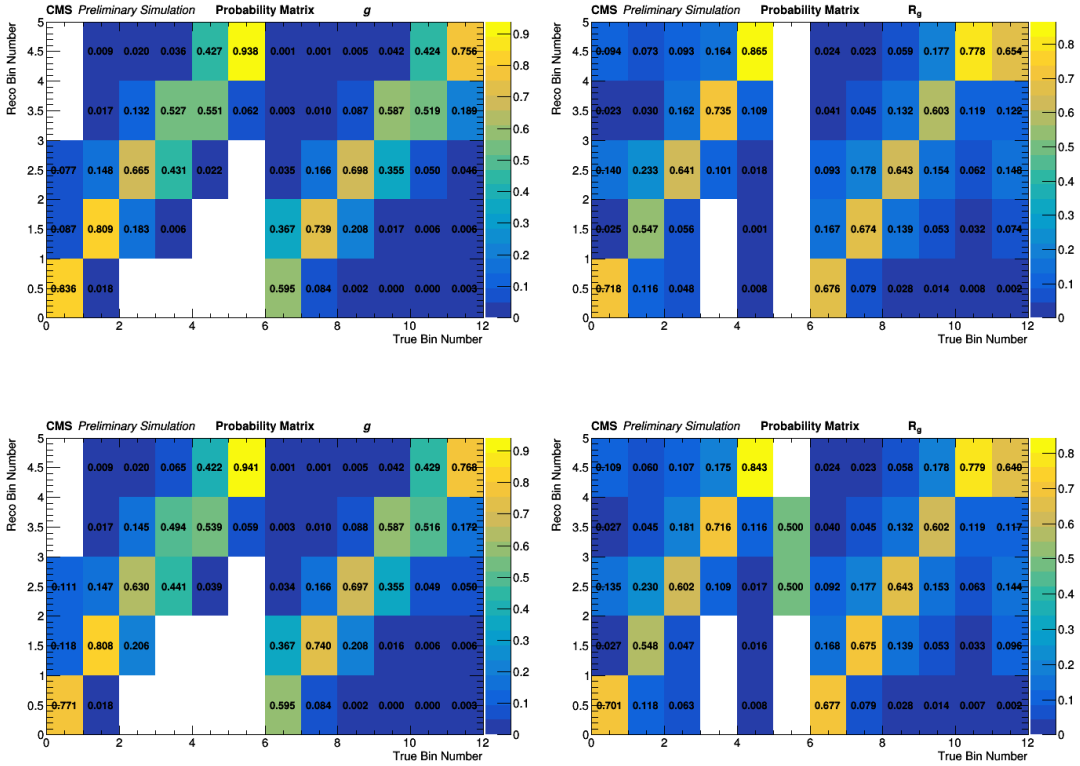


Figure 4.23: 2D Probability matrices for Girth (g) defined as (g True, g Reco) (left) and for R_g (right) defined as (R_g True, R_g Reco) for $x_{\gamma j} > 0.4$ (top) and $x_{\gamma j} > 0.4$ (bottom) in pp simulation in PYTHIA8. The numbers in the bins represent the normalized probability for each truth bin. The binning also includes the underflow/overflow bins to account for out-of-phase-space migrations. The first 6 bins in the x-axis correspond to the underflow in $x_{\gamma j}$.

diagonalities that are due to combinatorial subleading prongs selected by SoftDrop. The main difference is that, due to larger underlying event (UE) activity in PbPb collisions there is a broadening in the detector response for the substructure variables. The effect is more notable for girth(g), since it is a p_{T} weighted sum of the distances of the constituents, whereas R_g is more resilient to this effect.

In addition to smearing effects due to the larger underlying event activity in PbPb collisions, there is also the presence of residual fake prongs at detector level due to the UE which contribute to the off-diagonal element of the response matrix at large R_g values. These contributions are largely mitigated in our measurement with the $z_{cut} = 0.2$ and small $R = 0.2$ jets, and the residual amount is treated as a bin-to-bin migration in our unfolding. The remaining fraction of fake prongs is sufficiently small and can be dealt with the unfolding as we demonstrate with the closure tests.

4.8 Systematic uncertainties

Theoretical and experimental uncertainties are propagated in the unfolding procedure through variations of the response matrix, as well as the purity, and efficiency corrections. The following sources of systematic uncertainties are considered:

Physics model dependence: Regularized unfolding introduces a bias towards a MC generator-level spectrum in the unfolded solution to reduce the sensitivity to statistical fluctuations. To assess this bias, we use other assumptions for the prior spectrum to cover a reasonable range of possibilities. In addition, this results in a change in the detector migration matrix itself that is used for the unfolding corrections.

In pp collisions, the nominal MC sample used for unfolding is PYTHIA8 with the CP5 tune. We use HERWIG7.1.4 [179, 180] simulated events as an alternative MC sample, using the CH3 tune [169]. The HERWIG7 generator has an angular-ordered shower to account for color coherence effects, distinct from the p_{T}-ordered shower of PYTHIA8. HERWIG7 uses the cluster fragmentation model, different from the string model used in PYTHIA8. We take the difference of the unfolding determined with these variations relative to the nominal unfolded results as the respective systematic uncertainty. The resulting uncertainty is symmetrized bin by bin.

For PbPb collisions, we considered a variation of the PYTHIA8 CP5 events with a modification of its quark-gluon jet fraction, which is used as a proxy of medium-modification effects. It is expected that medium-induced jet modifications in general leads to a larger number of particles and momentum broadening. One way of assessing this effect in the corrections is by increasing the fraction of gluon jets. Quark and gluon jets have different shapes in R_{g} and g; quark jets tend to be narrower than gluon jets on average. The modification of the quark-gluon jet fraction in simulation is done by fitting the jet substructure observables at the detector level using a template fit with quark and gluon jet templates from simulated PYTHIA8 CP5 events. The template fitting yields a sample

of MC events with a larger fraction of gluon-initiated jets (from 45% to about 65%) for $x_{\gamma j} > 0.4$ while reducing the fraction of gluon-initiated jets (from 25% to about 12%) for $x_{\gamma j} > 0.8$ compared with the generator level yields, which we use to reweigh the nominal PYTHIA8 CP5 sample. The unfolding procedure is repeated with this variation of the PYTHIA8 CP5 sample using reweighted events according to the quark/gluon jet fraction found in the template fit. The symmetrized difference with respect to the nominal result is used as the respective systematic uncertainty.

Regularization bias uncertainty: The optimal number of iterations in D'Agostini unfolding has a dependence on the particle level prior spectrum used for regularization. To quantify the uncertainty associated with the prior spectrum choice, we use the optimal number of iterations obtained with MC sample variations described in the previous paragraph (quark/gluon fraction changes for PbPb, HERWIG7 for pp) and compare it with the results obtained with the nominal number of iterations.

PF candidate energy scale uncertainty: To assess the impact of the individual energy calibration uncertainties of the PF candidates that are used for the jet substructure, we scale the four-momenta of the charged hadron and photon PF candidates by $\pm 1\%$ and of neutral hadron PF candidates by $\pm 3\%$ at the detector level in simulation [21]. The PF energy scale variations are done in an uncorrelated way, i.e., an up or down shift in the four-momenta for each PF candidate species at a time, such that six different variations of PF energy scale shifts are evaluated. The unfolding corrections are repeated for each variation, and the difference with respect to the nominal unfolding result is evaluated. This has a different effect in R_g and g. For R_g, it changes the fraction of subjets that previously failed (or passed) the soft-drop grooming condition, and it mildly modifies the substructure distribution itself via smearing of the subjet p_T. For g, since it is defined as the sum of the products of p_T of the constituents and their y–ϕ distance to the jet axis, the shift has a direct effect on its shape and thus the effect is stronger than for R_g.

Jet energy scale (JES) uncertainties: The uncertainty in the JES is evaluated from dijet and γ + jetp_{T} balancing methods as discussed in Section 3.3.2 [160]. The uncertainty in the JES in pp collisions is around 3–4%, increasing as a function of $|\eta|$. In PbPb collisions, the uncertainty is around 4%, except for the barrel-endcap transition region ($1.3 < |\eta| < 1.6$), where it can become as large as 10%. An additional source of uncertainty in the JES is considered in PbPb collisions to take into account the differences in the particle mixture in simulation and data due to jet energy loss [84]. We construct an alternative set of unfolding corrections by varying the JES within its uncertainties. The JES uncertainty does not affect the jet substructure itself, since it is related to the calibration of the full jet. However, it can lead to migration effects via the $x_{\gamma \mathrm{j}} > 0.4$ (0.8) threshold used in our selection.

Jet energy resolution uncertainties: The uncertainty in the jet energy resolution is evaluated from a dijet balancing method as discussed in Section 3.3.2 [160]. In pp collisions, the uncertainty in the resolution is in the range of 2–4% in the barrel region, but is larger in the endcap and transition regions, where it varies in the range of 10–20%, depending on η. To propagate the uncertainty from jet p_{T} resolution, the data-to-simulation resolution scale factors are varied according to their uncertainties and the unfolding corrections are repeated. In PbPb collisions, there is an additional contribution to the jet energy resolution uncertainty because of the modeling of the UE in simulation with HYDJET. This uncertainty is evaluated by comparing the energy in randomly-distributed cones in data and simulation. The difference in data and simulation on the energy distributions obtained using the random-cone method is used to estimate the effect on the jet resolution. To estimate this uncertainty, we shift the centrality interval in simulation. To account for these data-to-simulation differences, in practice the centrality in simulation is shifted by 4.5% for the nominal results, i.e, 0–30% central events in data correspond to 4.5–34.5% in simulation. For the respective uncertainty, we consider the intervals of 3–33% and 6–36% for the down and up variations, respectively.

Photon background subtraction uncertainty: Nominally, the normalization of the photon background from neutral hadron decays is determined using a template fit of the $\sigma_{\eta\eta}$ distribution using simulated signal photon+jet events and the shower shape distribution in data in a region dominated by background (sideband in $\mathcal{I}$). As an alternative method, we calculate the normalization of the background using control samples in data via the ABCD method, as explained in Section 4.4. The unfolding is repeated after this neutral-hadron photon decay background subtraction variation and the corresponding difference with respect to the yields extracted with the template fit is considered as the systematic uncertainty, which is symmetrized.

Response matrix statistical uncertainties: We consider the statistical uncertainties in the MC sample used to construct the response matrix as an additional systematic uncertainty, which is propagated in the unfolding procedure at each iteration.

The uncertainties are added in quadrature bin-by-bin for the final combination of systematic uncertainties. The uncertainties are considered to be bin-to-bin fully correlated, except for the statistical uncertainties of the unfolded distributions and the response matrix statistical uncertainties, whose covariance matrices are determined directly as part of the unfolding procedure ROOUNFOLD. The relative bin-by-bin uncertainties are shown in Table. 4.6 for $x_{\gamma j} > 0.4$ and in Table.4.7 for $x_{\gamma j} > 0.8$ in both pp and PbPb collisions. The dominant contribution to the measurement uncertainty is the model dependence used in the corrections, with the PF energy scale uncertainty being the subleading one. The uncertainties grow at large values of g and R_g. Statistical uncertainties are of the same order as the systematic uncertainties for the PbPb measurement, whereas they are smaller than the systematic ones for pp. Since the radiation pattern is modified in nontrivial ways in the case of PbPb collisions and since the PbPb and pp data were collected in different years, we consider the systematic uncertainties to be uncorrelated between the PbPb and pp measurements.

Table 4.6: Summary of the minimum and maximum bin-by-bin percentual relative uncertainties for $x_{\gamma j} > 0.4$.

Uncertainty source	g		R_g	
	pp	PbPb	pp	PbPb
Physics model dependence	1.3–7.5	1.2–2.5	0.2–5.3	0.9–5.4
Regularization bias	0.1–0.7	$\lesssim$0.1–1.2	$\lesssim$0.1	$\lesssim$0.1
Photon PF energy scale	0.4–1.5	1.8–5.1	0.2–0.6	0.1–2.9
Charged hadron PF energy scale	0.6–4.1	0.5–5.3	$\lesssim$0.1–0.5	$\lesssim$0.1–1.6
Neutral hadron PF energy scale	0.1–1.5	0.2–5.3	0.1–0.5	0.3–2.4
JES	0.2–3.3	0.2–2.6	0.1–1.5	0.3–3.3
JER	$\lesssim$0.1–2.0	$\lesssim$0.1–3.7	$\lesssim$0.1–0.2	0.1–1.8
Centrality	—	0.6–4.0	—	$\lesssim$0.1–2.4
Photon background subtraction	0.1–0.3	$\lesssim$0.1	$\lesssim$0.1–0.2	$\lesssim$0.1
Response matrix statistical	1.0–2.9	1.4–4.5	0.9–2.2	1.4–3.6
Total systematic	2.2–9.8	2.7–10.7	1.3–6.0	2.7–8.5
Total statistical	1.4–3.5	3.5–7.6	1.4–2.5	3.6–6.4

Table 4.7: Summary of the minimum and maximum bin-by-bin percentual relative uncertainties for $x_{\gamma j} > 0.8$.

Uncertainty source	g		R_g	
	pp	PbPb	pp	PbPb
Physics model dependence	0.3–8.9	0.1–7.3	0.6–3.0	0.1–5.7
Regularization bias	0.3–0.9	0.1–2.0	$\lesssim$0.1–1.6	0.6–2.9
Photon PF energy scale	0.1–1.3	0.1–4.4	0.1–1.4	0.2–0.9
Charged hadron PF energy scale	0.5–3.7	0.1–8.6	$\lesssim$0.1–2.5	0.1–1.3
Neutral hadron PF energy scale	0.0–2.4	0.2–7.5	0.1–1.9	0.1–3.7
JES	0.7–4.8	1.5–7.8	0.0–5.3	2.0–6.8
JER	$\lesssim$0.1–1.0	0.2–3.0	$\lesssim$0.1–2.1	0.3–3.0
Centrality	—	0.4–5.9	—	0.1–2.5
Photon background subtraction	$\lesssim$0.1–0.1	$\lesssim$0.1–0.1	0.1–0.1	$\lesssim$0.1–0.1
Response matrix statistical	1.0–3.8	1.5–5.9	1.0–3.2	1.1–4.0
Total systematic	1.6–11.7	2.5–14.8	2.7–7.5	2.6–10.6
Total statistical	1.4–5.1	4.2–14.6	1.6–3.9	3.6–11.6

The detailed bin-by-bin variation of the relative uncertainties for jet girth g and groomed jet radius R_g are shown in Figs. 4.24, 4.25, 4.26 and 4.27. Figure 4.24 shows the uncertainty breakdown in PbPb collisions for $x_{\gamma j} > 0.4$, where the statistical uncertainty in data is commensurate of the total systematic uncertainty. The uncertainties are larger in the tails of the distribution with the leading uncertainty for R_g being the fragmentation modeling and the various PF scale uncertainties dominating the g uncertainties. Figure 4.25 shows similar trends in PbPb collisions for $x_{\gamma j} > 0.8$. Due to the tighter event selection the statistical uncertainty is larger in all bins, also impacting the other systematic variations. Figures 4.26 and 4.27 are the relative uncertainties in pp collisions for $x_{\gamma j} > 0.4$ and $x_{\gamma j} > 0.8$, respectively. The overall uncertainty is smaller in pp collisions with respect to PbPb and follow similar trends due to a better MC modeling of the data and higher event yields in the analysis for the g and R_g observables. A significant difference is observed for the fragmentation modeling uncertainty for the SoftDrop untagged events (shown in the negative [-0.05,0] bin) caused due to the differences in PYTHIA8 and HERWIG7 used to estimate it unlike in PbPb where the PYTHIA8 MC is reweighted to estimate this uncertainty. The observables considered in the analysis are not sensitive to the integrated luminosity uncertainty, since they are normalized to the total number of selected events. The photon energy scale uncertainties have a negligible effect on the measured jet substructure distributions.

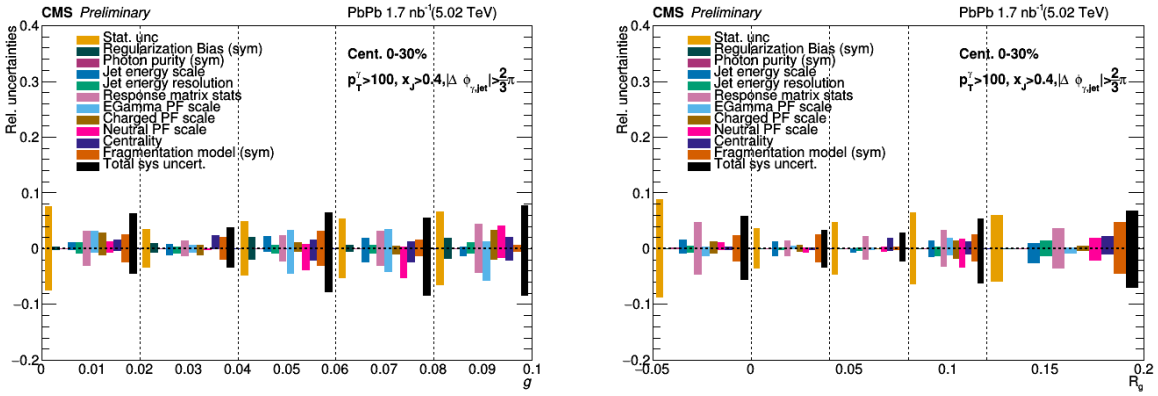


Figure 4.24: Systematic uncertainties for jet girth g (left) and groomed jet radius R_g (right) of jets recoiling from photons in PbPb collisions for $x_{\gamma j} > 0.4$.

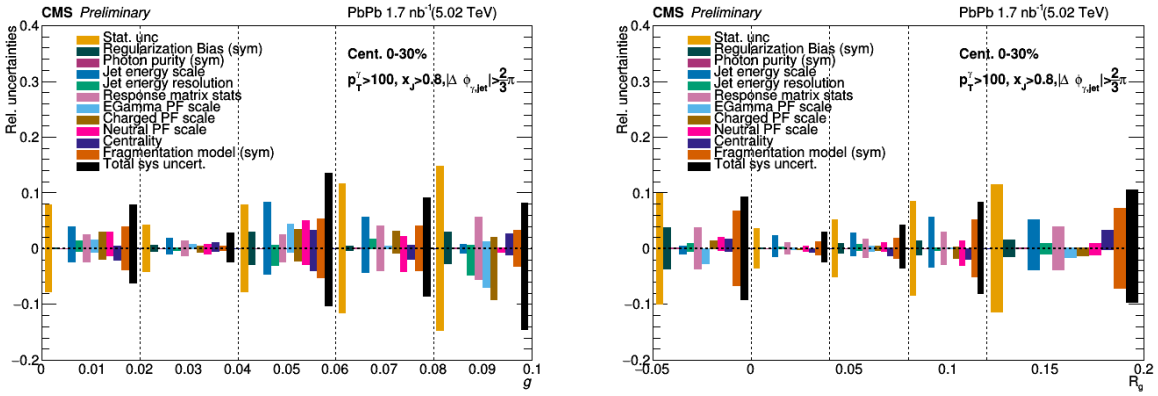


Figure 4.25: Systematic uncertainties for jet girth g (left) and groomed jet radius R_g (right) of jets recoiling from photons in PbPb collisions for $x_{\gamma j} > 0.8$.

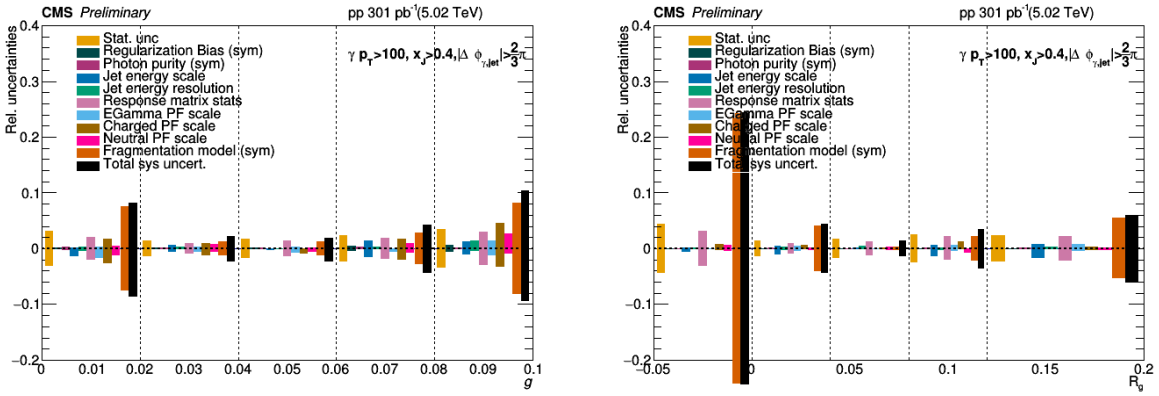


Figure 4.26: Systematic uncertainties for jet girth g (left) and groomed jet radius R_g (right) of jets recoiling from photons in pp collisions for $x_{\gamma \mathrm{j}} > 0.4$.

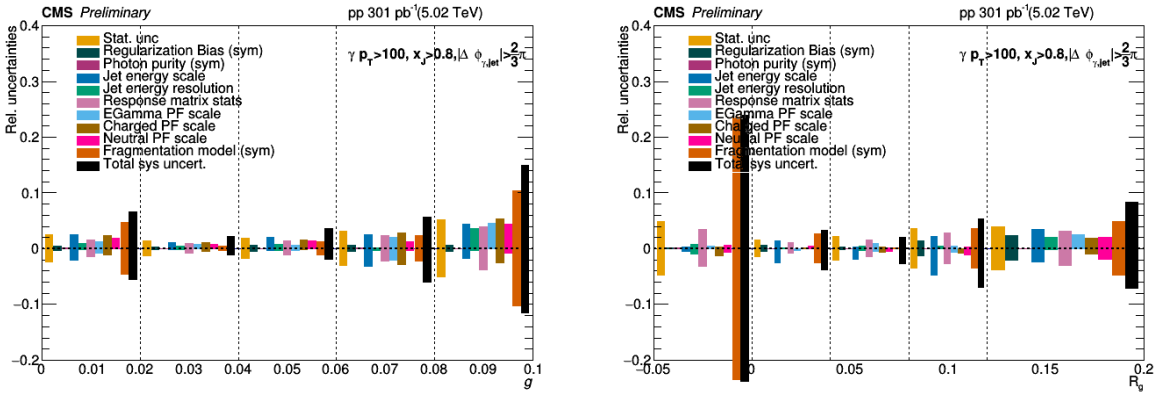


Figure 4.27: Systematic uncertainties for jet girth g (left) and groomed jet radius R_g (right) of jets recoiling from photons in pp collisions for $x_{\gamma \mathrm{j}} > 0.8$.

Chapter 5

Cross checks and validation

5.1 Event reweighting checks

5.1.1 Jet p_{T} reweighting

The jet p_T spectrum in central PbPb is reweighted at detector level to match the data distribution as shown in Fig.5.1.

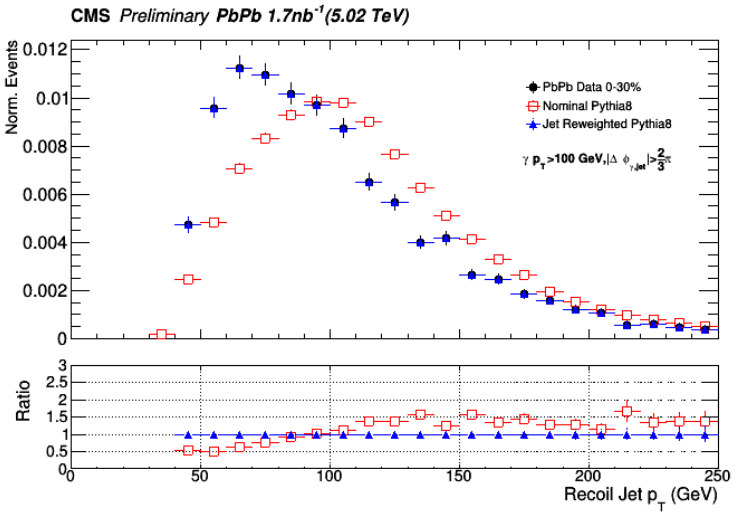


Figure 5.1: Jet p_T Spectrum for 0-30% Central PbPb in Data, Nominal PYTHIA8 Embedded and Reweighted PYTHIA8

The substructure (R_g and g) are unfolded using this reweighted response. The result is statistically consistent with the reported nominal result as shown in Fig.5.2.

This should be regarded as a check, and not as an additional systematic uncertainty in our measurement. The unfolded distribution extracted using the jet p_T-reweighted MC sample is already covered by the combination in quadrature of the systematic uncertainties and statistical uncertainties, so this already satisfies the check.

Another reason this should not be added as a systematic uncertainty is that the reweighting of MC to data of the jet p_T spectra enhances the weight of events where the recoil jet is the result of a large angle parton branching, and not a jet initiated by a quark or gluon from the hard scattering process. Previous experience from inclusive jet analyses have shown that quenched jets look more like the jets initiated by the partons from the hard scattering than jets that are generated from large angle splittings in the shower. It would be, in a way, promoting "anomalous" small R jets that result from the parton shower, rather than primary hard jets.

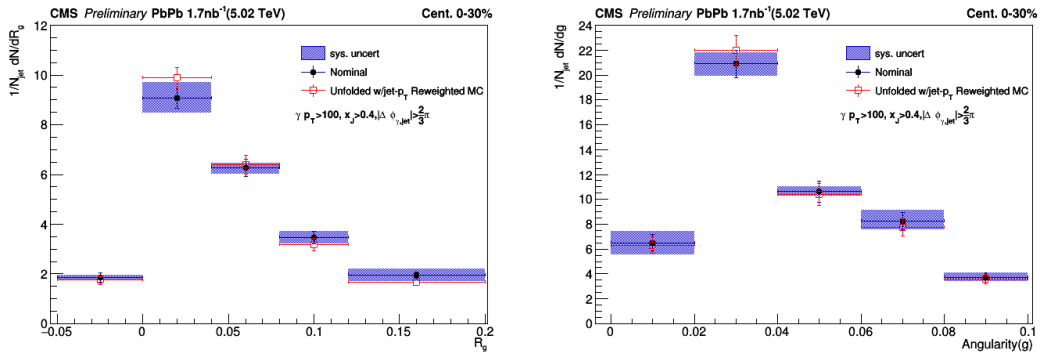


Figure 5.2: Unfolded R_g (left) and g (right) comparison before and after reweighting the jet p_T spectrum embedded PYTHIA8 with 0-30% central PbPb data

5.1.2 Event density reweighting (ρ)

The event density is defined by the observable ρ as defined in equation 3.9 in Sec. 3.3.1. It is a global observable useful in the determination of the centrality of a given event. The ρ spectrum in central PbPb obtained from MC is reweighted at detector level to match the data distribution as shown in Fig.5.3.

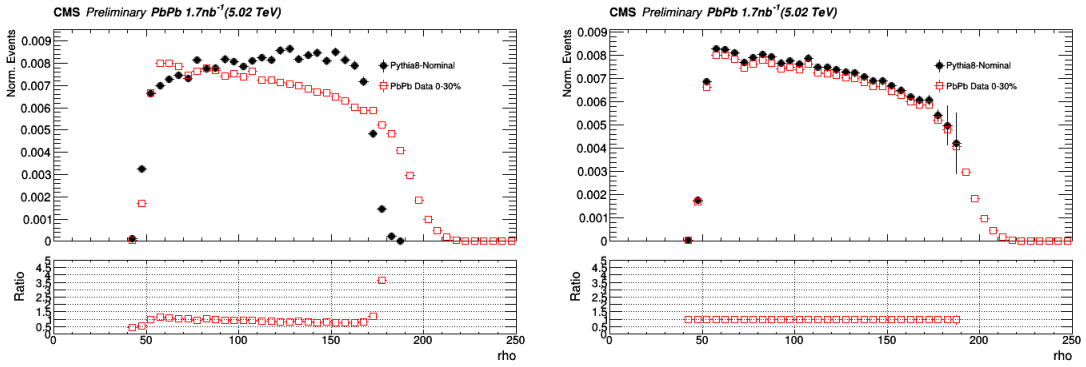


Figure 5.3: Rho Distribution in MC before and after reweighting compared to PbPb Data

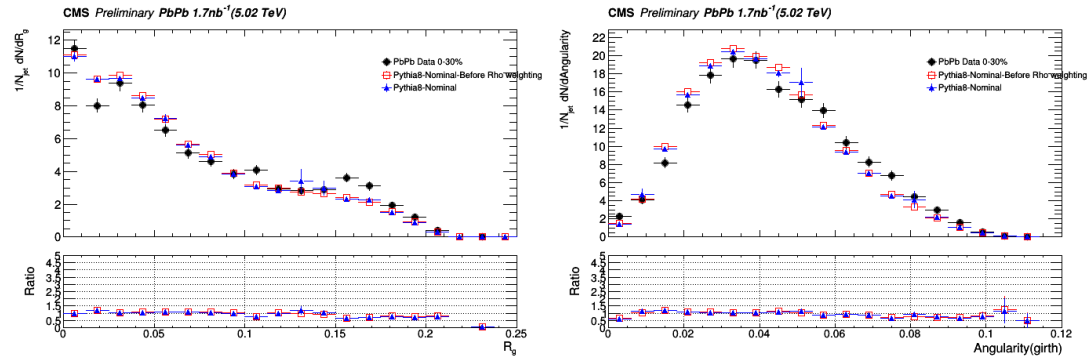


Figure 5.4: Detector level R_g (left) and g (right) comparison before and after reweighting the ρ spectrum

The effect of the ρ reweighting on the substructure variables (R_g and g) is negligible as shown in Fig. 5.4

5.1.3 Vertex (v_z) reweighting

The v_z distribution in central PbPb in MC is reweighted at detector level to match the data distribution as shown in Fig.5.5.

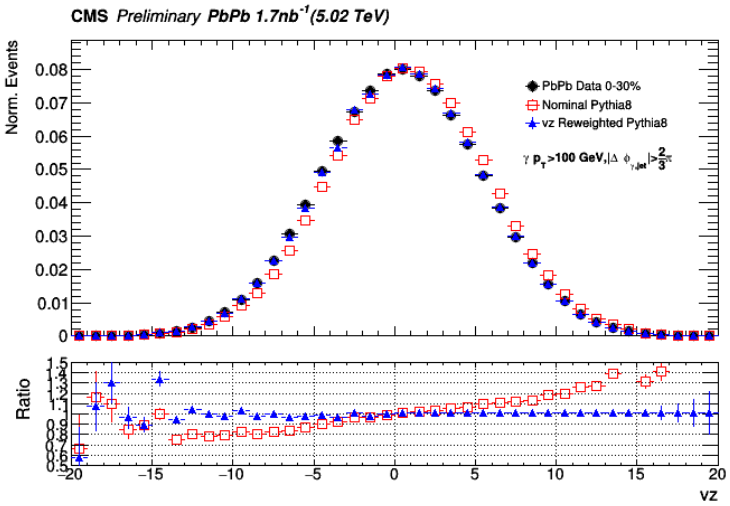


Figure 5.5: Comparison of vertex v_z distribution for 0-30% central PbPb in Data, Nominal PYTHIA8 Embedded and Reweighted PYTHIA8

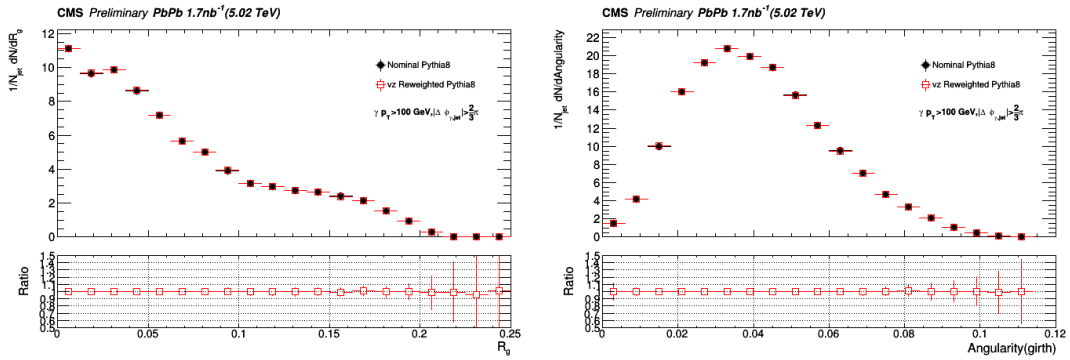


Figure 5.6: Detector level R_g (left) and g (right) comparison before and after reweighting the vz spectrum

The effect of the v_z reweighting on the substructure variables (R_g and g) is negligible as shown in Fig. 5.6

5.2 Residual jet energy scale

The jet energy corrections used in this analysis were derived for inclusive dijet samples as discussed in Section 3.3.2. The corrections are dependent on the quark/gluon content of the jets which are different in photon-tagged jet samples and inclusive jet samples. This can lead to a residual jet energy scale and the jet p_{T} and η dependence of this residual is verified. Figures 5.7 and 5.8 show the residuals for pp and central (0-30%) PbPb MC samples. They are within the expected resolution and addressed by the jet systematic uncertainties.

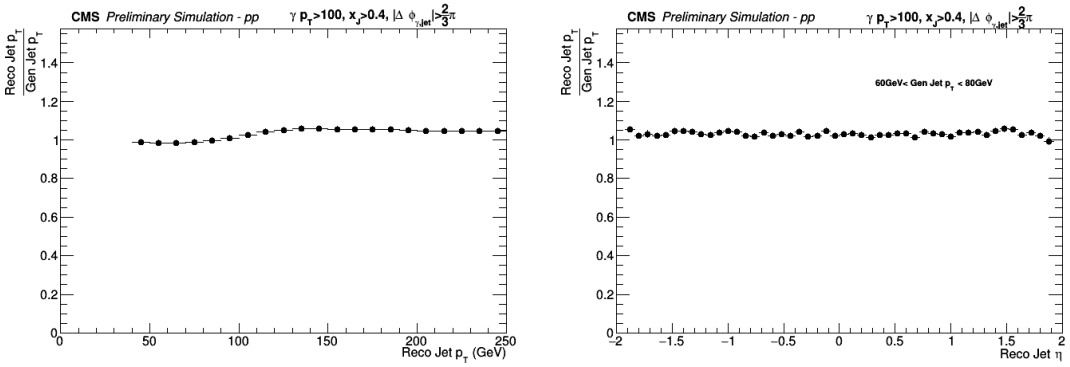


Figure 5.7: Residual jet energy scale in the pp PYTHIA8 sample, shown as a function of jet p_T(Left) and jet η (Right) .

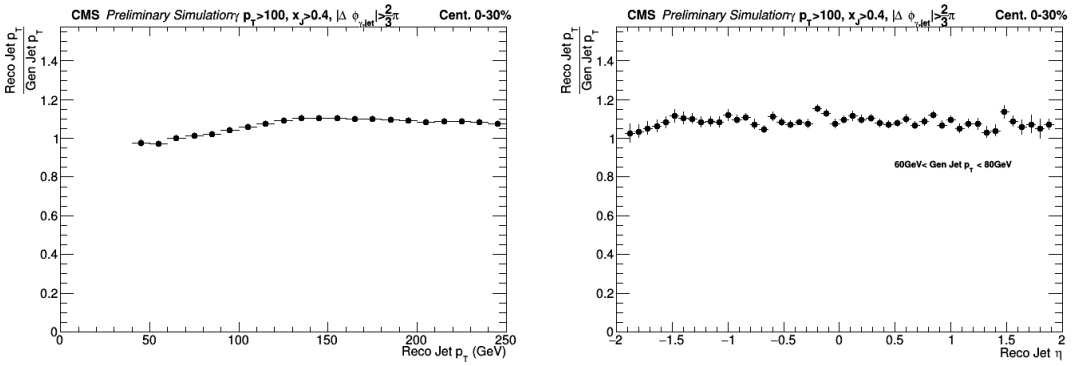


Figure 5.8: Residual jet energy scale in the central 0-30% PbPb embedded PYTHIA8 sample, shown as a function of jet p_T(Left) and jet η (Right) .

5.3 Residual combinatorial background estimation

The shapes of the $\Delta\phi_{\gamma j}$ as shown in Fig. 5.9 distributions in PbPb and pp collisions are compatible across the entire range, which suggests that the residual combinatorial jet background from the UE in PbPb collisions is negligible within the statistical sensitivity of the measurement.

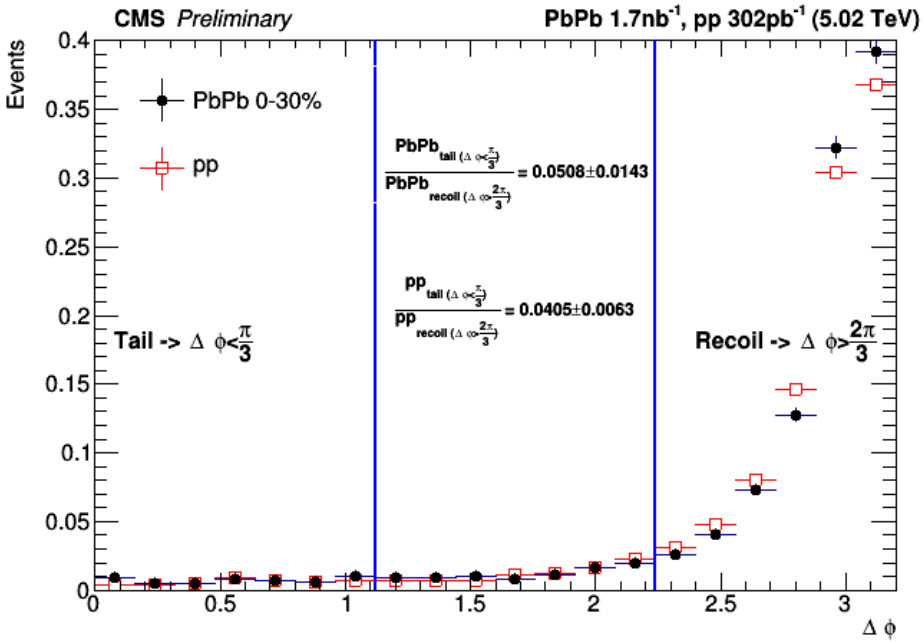


Figure 5.9: Raw-level comparisons between pp and PbPb $0-30\%$ central $\Delta\phi_{\gamma\text{jet}}$ distributions. The conclusion here is that the same angular correlations between the photon and the jets is found in pp and PbPb $0-30\%$ central collisions, meaning that the combinatorial background at the jet-level is negligible.

5.4 Trigger efficiency

The L1 and HLT trigger efficiency was checked in central PbPb Minimum Bias data. As expected from previous studies discussed in Section 3.2, the triggers are fully efficient for reconstructed photons of $p_{\mathrm{T}}^{\gamma} > 100\,\mathrm{GeV}$ as shown in Fig. 5.10.

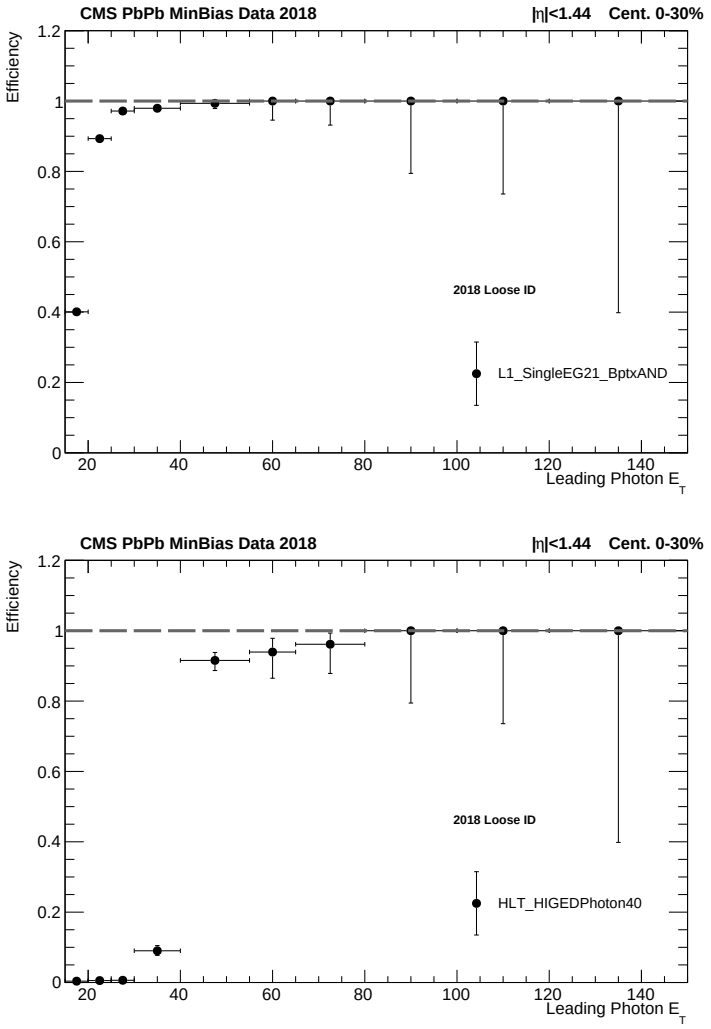


Figure 5.10: The trigger efficiency at L1 for a threshold of 21 GeV (upper) and corresponding HLT for a threshold of 40 GeV (lower) as a function of the leading offline photon E_T used in the analysis in $0-30\%$ central PbPb Minimum Bias data collected in 2018. The trigger is fully efficient for photons above 100 GeV.

5.5 Substructure residual

In order to inspect in detail the response matrix, we consider the residual distributions. The residual distribution of R_g is defined as $\frac{R_g^{reco-level} - R_g^{true-level}}{R_g^{true-level}}$. It is defined similarly for the jet p_T and g. The residuals for Recoil Jet p_T, Girth (g) and R_g for $60 < p_{\mathrm{T,jet}}^{\mathrm{true}} < 80\,\mathrm{GeV}$ are shown in Fig.5.11. We see that the background shifts the jet p_T residuals distribution to higher reconstructed values and that it increases the width of the distribution as compared to solely detector effects as expected. These residuals provide a one dimensional representation of the response matrix used in unfolding the detector level distribution to their true level. Large shifts and asymmetry in the residual distribution is a sign of large bin to bin migrations.

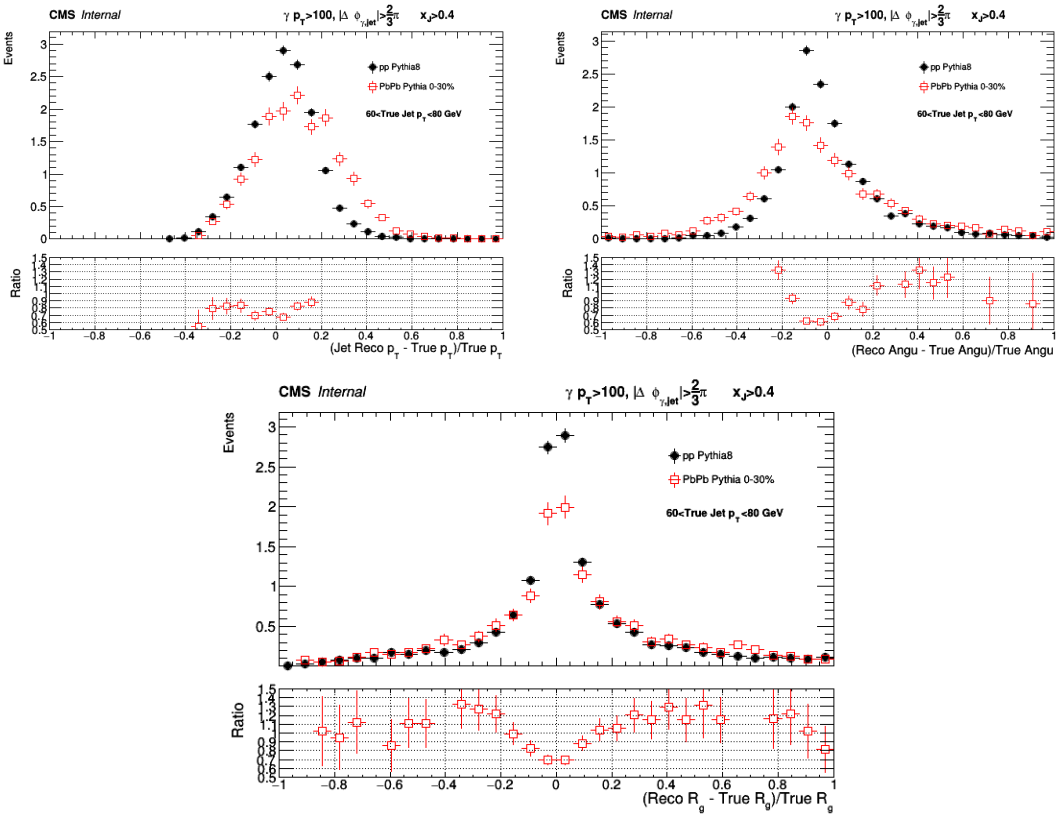


Figure 5.11: Residual distribution for Recoil Jet p_T (upper left), Girth (g) (upper right) and R_g (lower) in pp PYTHIA8 compared to $0-30\%$ PbPb. The asymmetric shift in Recoil Jet p_T and Girth (g) is likely due to reconstruction inefficiencies at detector-level.

5.6 Recoil jet selection

The signal processes of interest in the study of photon tagged jet substructure as discussed in Section 3.2 at leading order result in a single parton producing the jet shower. However, the jet clustering algorithm might result in the decay products being clustered into multiple jets due to factors such as changes in the energy and momentum due to the UE. We compared the study of the leading recoil jet and all recoil jets in the cone of $\Delta\varphi_{\gamma j} > \frac{2}{3}\pi$ as shown in Fig. 4.5. The detector level comparison of the $x_{\gamma \mathrm{j}}$ distribution as well as the substructure distributions g and R_{g} are shown in Figs. 5.12 and 5.13. We see that in 95% of the events in central PbPb data there is only one jet in the recoil cone. The results of this analysis are yields normalized to the number of jets satisfying the kinematic selection. Solely considering the leading recoil jet, the complexity of correlations between the jets in the same event are removed. Thus, we consider only the leading recoil jet as our selected jet for this analysis as the difference between them can be considered negligible.

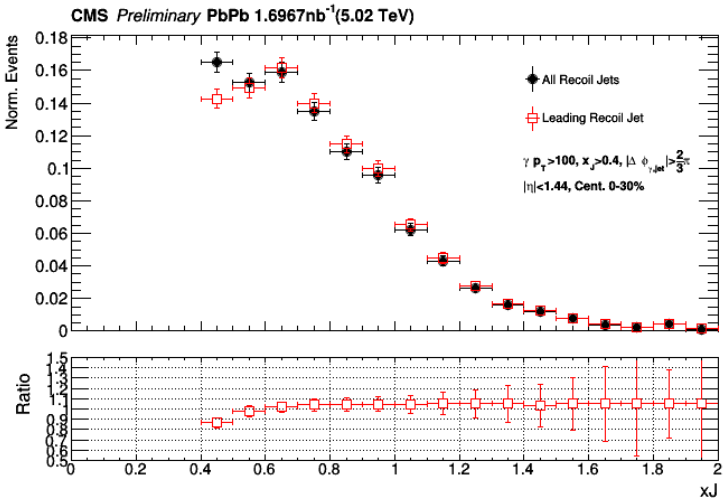


Figure 5.12: Distribution of $x_{\gamma \mathrm{j}}$ in central 0-30% PbPb between leading recoil jet and all recoil jets. The difference between considering all recoil jets and leading recoil jet is negligible.

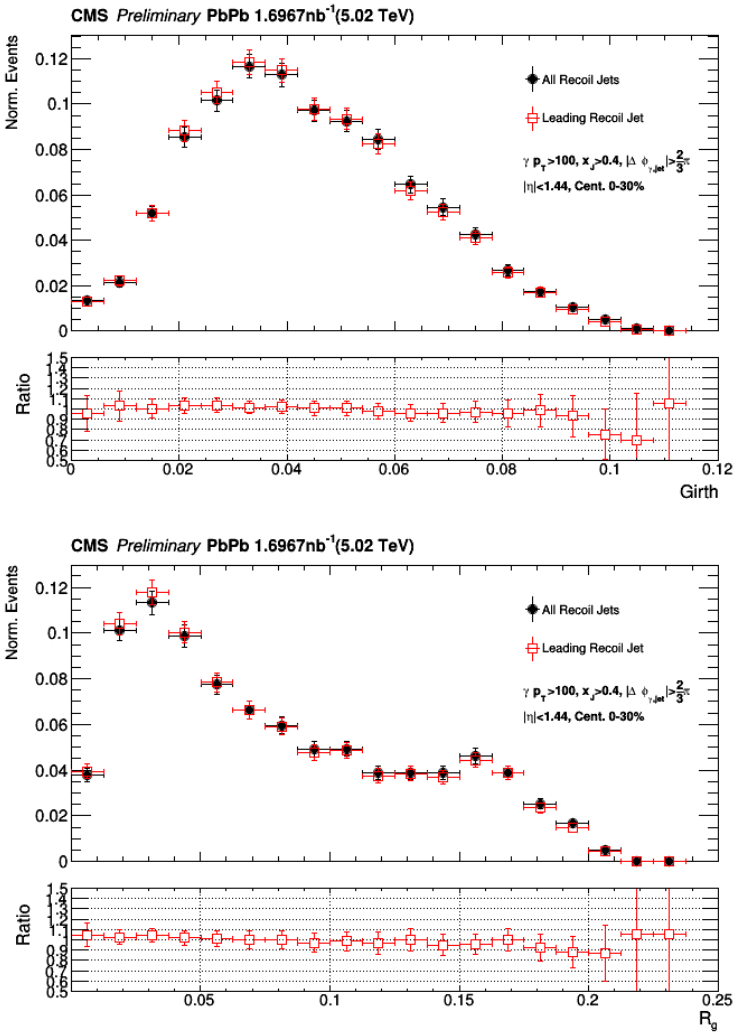


Figure 5.13: Jet substructure distributions of girth g (left) and R_g (right) in central 0-30% PbPb between leading recoil jet and all recoil jets. The difference between considering all recoil jets and leading recoil jet is negligible.

5.7 Quark gluon fraction fitting

For PbPb collisions, we consider that the dominant effect that should be considered for the prior uncertainty is due to jet quenching. The embedded PYTHIA8 simulated events that are used for the corrections do not account for jet quenching effects. We investigated the possibility of using Pyquen simulated events (which accounts for jet quenching effects in an ad hoc way) as an alternative prior model in order to account for these effects. However, we found that the jet substructure is severely modified in that case in a way that is not consistent with what we observe in the data. This can lead to an overestimation of the modeling uncertainties associated to the unsmearing procedure used in this analysis. Since the predictions do not reasonably describe the data, we considered an alternative: a variation of the PYTHIA8 simulated events with a different quark-gluon fraction as a proxy for a quenched Monte Carlo. This is under the assumption that the inclusion of jet quenching effects in the predictions should lead to a better description of the data.

The quark-gluon fraction is fitted in order to provide a better representation of the jet substructure at detector-level using a template fit. The jet flavor is assigned by the flavor association tagging available in CMSSW, as implemented in HiForest. Based on color factors, one expects quark jets to be narrower and have a harder fragmentation than gluon jets for the substructure observables considered in this note. Indeed, this is reflected in the R_g and girth g distributions. The difference in shapes for the quarks and gluons gives us a handle to do a robust template fitting procedure, since there is no degeneracy in the shapes. The template fits are done separately for the R_g and girth distributions. This variation of the PYTHIA8 embedded sample is used as an alternative MC prior, which accounts for jet quenching effects (or other missing higher order corrections) in an effective way. What we find with this procedure is that the data is better described with a slightly larger fraction of gluon jets than what is present in

the nominal sample. This is observed for both the R_g and girth distributions.

The original quark fraction is 64.6% for $x_{\gamma j} > 0.4$ and 74.4% for $x_{\gamma j} > 0.8$ obtained from particle level PYTHIA8 MC. The results of the template fits used to extract the quark/gluon fraction are shown in Fig. 5.14 and 5.15 for $x_{\gamma j} > 0.4$ and $x_{\gamma j} > 0.8$. As seen in those plots, the quark fraction is reduced down to 54% for $x_{\gamma j} > 0.4$ in order to better fit the data, i.e., the gluon fraction is increased, while for $x_{\gamma j} > 0.8$ the quark fraction is increased to 90%. In the same plots, one can see that the quark jet distribution and the gluon jet distributions are very different from one another; the R_g of gluon jets is broader, whereas the one of quark jets is narrower. This is consistent with the expectation that $x_{\gamma j} > 0.8$ selects harder fragmentation jets and is preferentially populated by narrow early vacuum jets originating primarily through quarks. Figure 5.16 shows the comparison of the results of unfolding the data with the extreme cases of considering purely quarks or purely gluons.

A reduction of the quark fraction is also seen for the inclusive $x_{\gamma j} > 0.4$ selection in pp as shown in Fig. 5.17 which is consistent with PbPb.

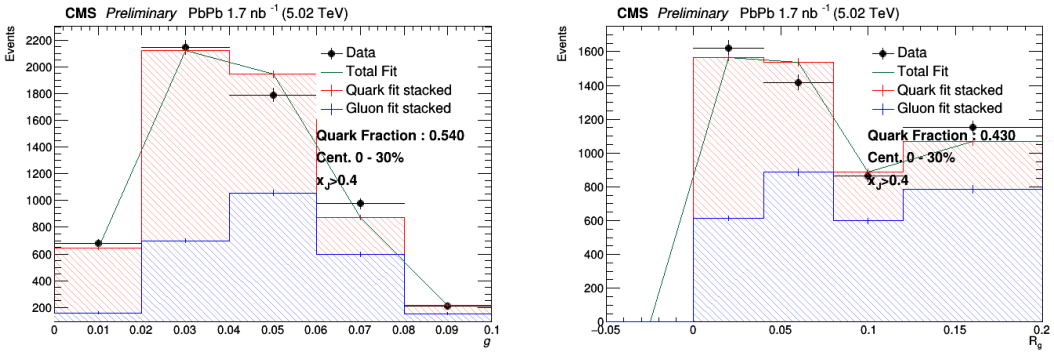


Figure 5.14: Template Fit of the g (left) and R_g (right) distribution at detector level using recoil quark jets and gluon jets for 0-30% PbPb for $x_{\gamma j} > 0.4$. The black points are the PbPb data with quark and gluon distributions from MC in red and blue stacked histograms

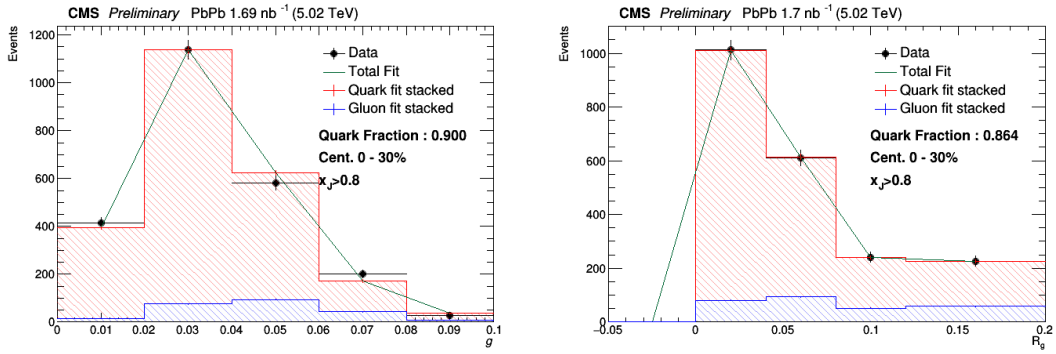


Figure 5.15: Template Fit of the g (left) and R_g (right) distribution at detector level using recoil quark jets and gluon jets for 0-30% PbPb for $x_{\gamma j} > 0.8$. The black points are the PbPb data with quark and gluon distributions from MC in red and blue stacked histograms

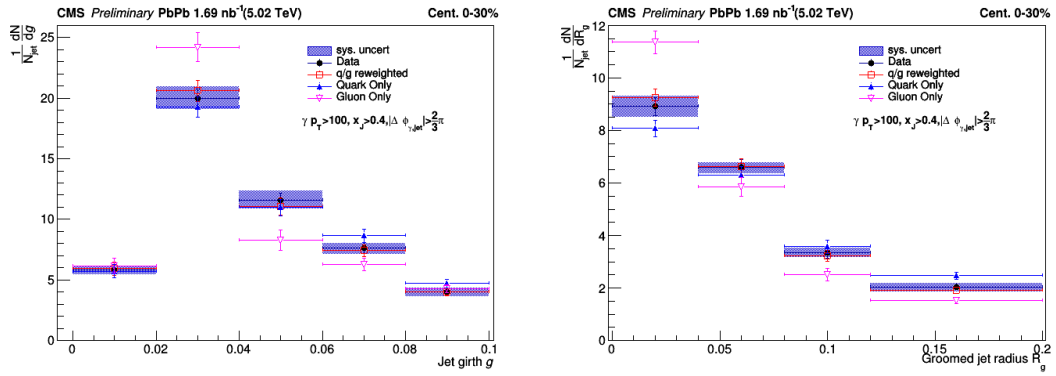


Figure 5.16: Result of unfolding with corrections derived from quark-only jets and gluon-only jets (PYTHIA8+HYDJET) for g and R_g. The vertical bars represent statistical uncertainties, the band represents the systematic uncertainties only

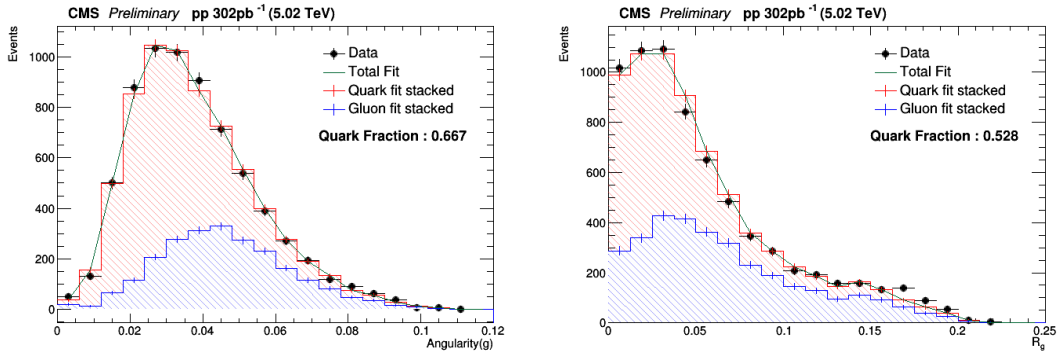


Figure 5.17: Template Fit of the g and R_g distribution at detector level using recoil quark jets and gluon jets for pp for $x_{\gamma j} > 0.4$.

5.8 Particle flow energy scale uncertainty bounds

The jet energy scale is derived by treating the jet as a whole object, not at the per-PF candidate level. The PF scale uncertainties quantify the effect at the substructure level, which is not propagated from the full JES uncertainty. The effect of the PF candidate energy scale shift is large at the substructure level, whereas it generates a shift on the full jet pT that is smaller than the full JES uncertainty. This approach has been followed in other recent CMS jet substructure analyses as well [21, 181].

The individual variations are part of the overall JES uncertainty, so in principle there is a small component of the JES that are indeed double counted. However, the JES uncertainty is not dominant, so the overall size of the uncertainties is not expected to reduce in the process of removing this component. The procedure to decompose the components in an attempt to remove the double counting is not trivial and not considered required in the scope of this analysis.

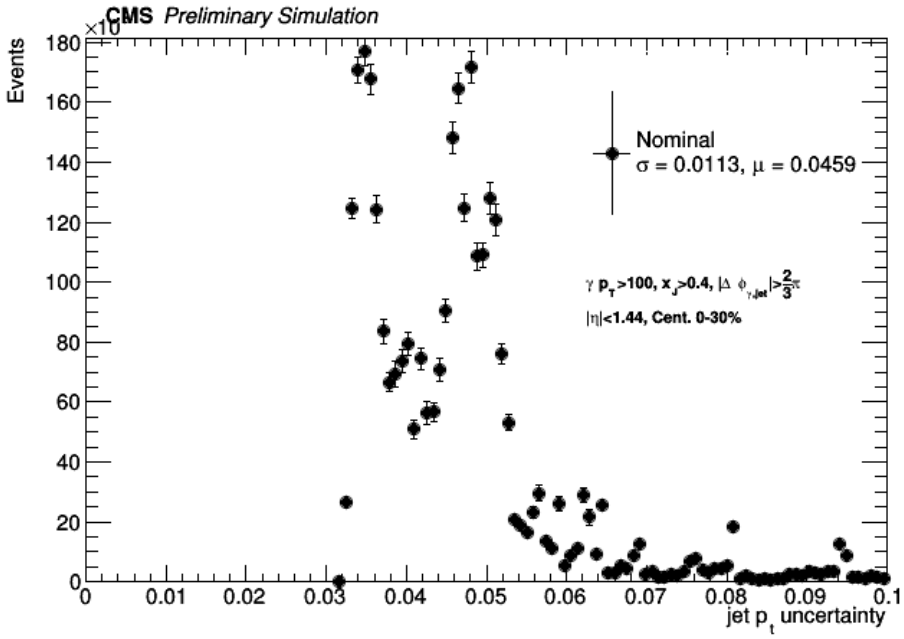


Figure 5.18: Jet energy scale uncertainty

Figure 5.18 shows the expected jet energy scale uncertainty in PbPb, with values in the ballpark of 3–5%. The impact of the individual PF scale uncertainties are estimated by considering the ratio of the reclustered CA jet p_{T} for the given PF scale shift with respect to the nominal jet p_{T} and comparing the variance of the distributions to the JES uncertainty. The individual PF scale uncertainty variations are shown in Fig. 5.19 to be of an order $< 0.5\%$ with the neutral PF having the largest effect.

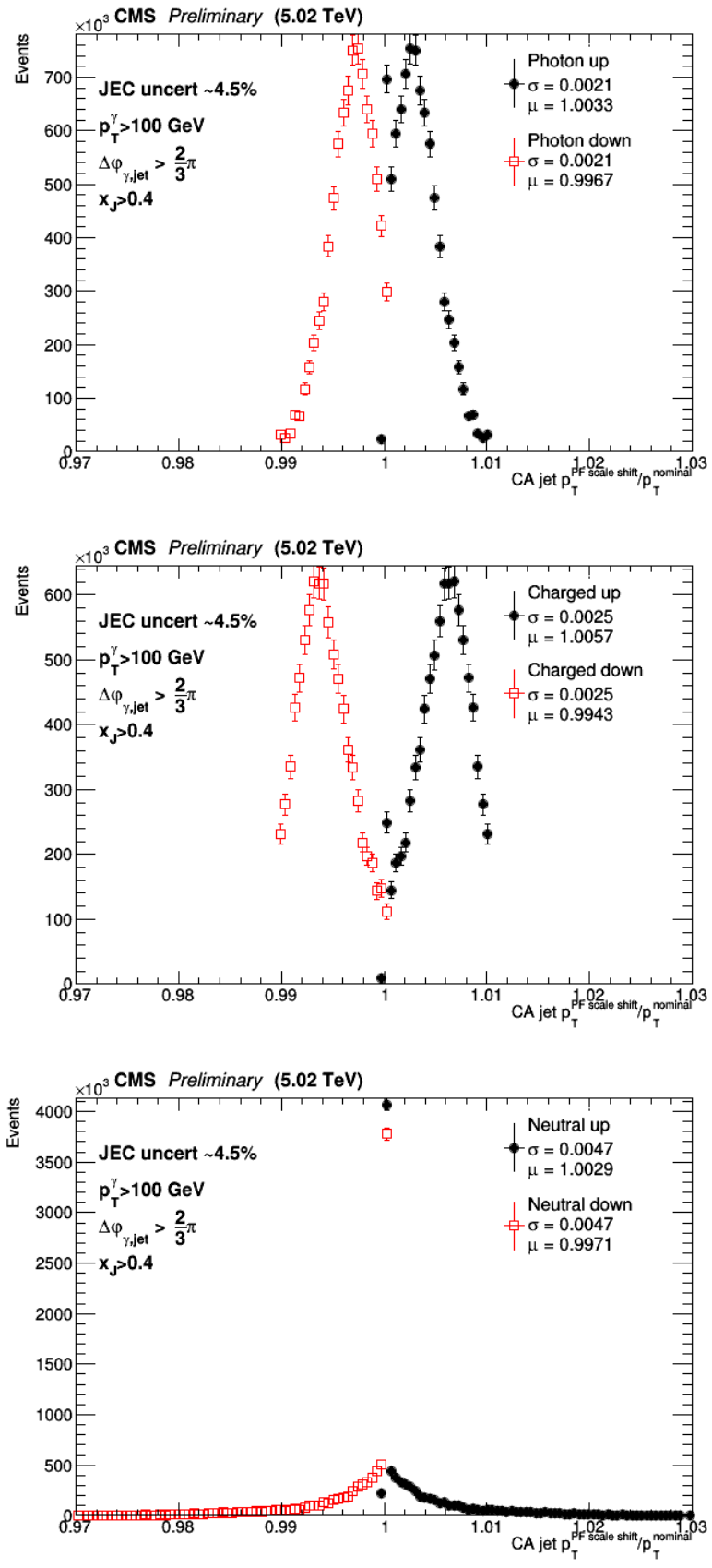


Figure 5.19: Individual PF scale variation compared to jet energy scale uncertainty

5.9 Correlations in systematic uncertainty

This thesis presents the analysis results for the ratio of the normalized yields of PbPb to pp for the jet substructure observables in Chapter 6. The PbPb data was collected in 2018 and the reference pp data was collected in 2017 with significant changes to the detector conditions. The various systematic uncertainties in this analysis are considered to be uncorrelated between the PbPb and pp systems, thus the uncertainty on the ratios are assumed to be fully uncorrelated for the results. Under ideal conditions, systematic uncertainties such as the PF constituent energy scales, JES, JER and photon background subtraction could have correlations between the two systems that can be cancelled in the ratio. A check was performed to estimate the maximal cancellations by considering these uncertainties to be fully correlated instead and compared to the fully uncorrelated uncertainties. Figures 5.20 and 5.21 show the comparison of the uncertainties for $x_{\gamma j} > 0.4$ and $x_{\gamma j} > 0.8$ respectively showing a negligible difference to the total uncertainty between the two cases. The leading uncertainties are not cancelled and the most optimistic fully correlated results do not show a significant difference, thus leaving the comparison as a check.

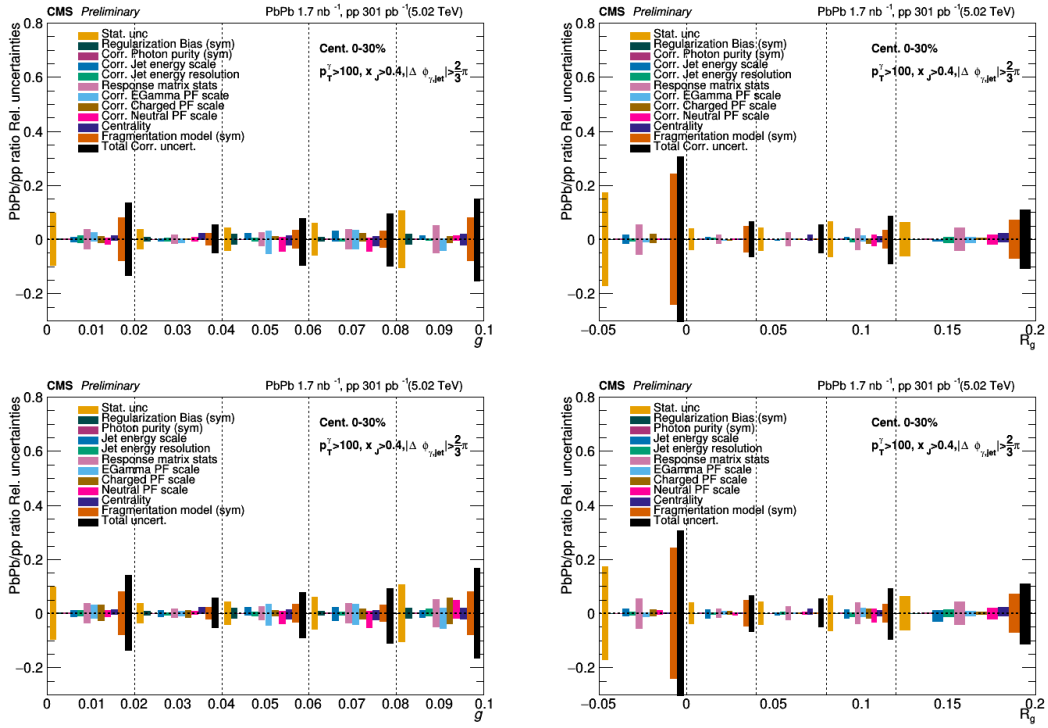


Figure 5.20: Systematic uncertainty breakdown for g (left) and R_g (right) of jets recoiling from photons in the ratio of PbPb to pp collisions for $x_{\gamma j} > 0.4$. A check was performed considering some of the systematic uncertainties as considered fully correlated(upper) and compared to the nominal results where no correlations are considered (lower).

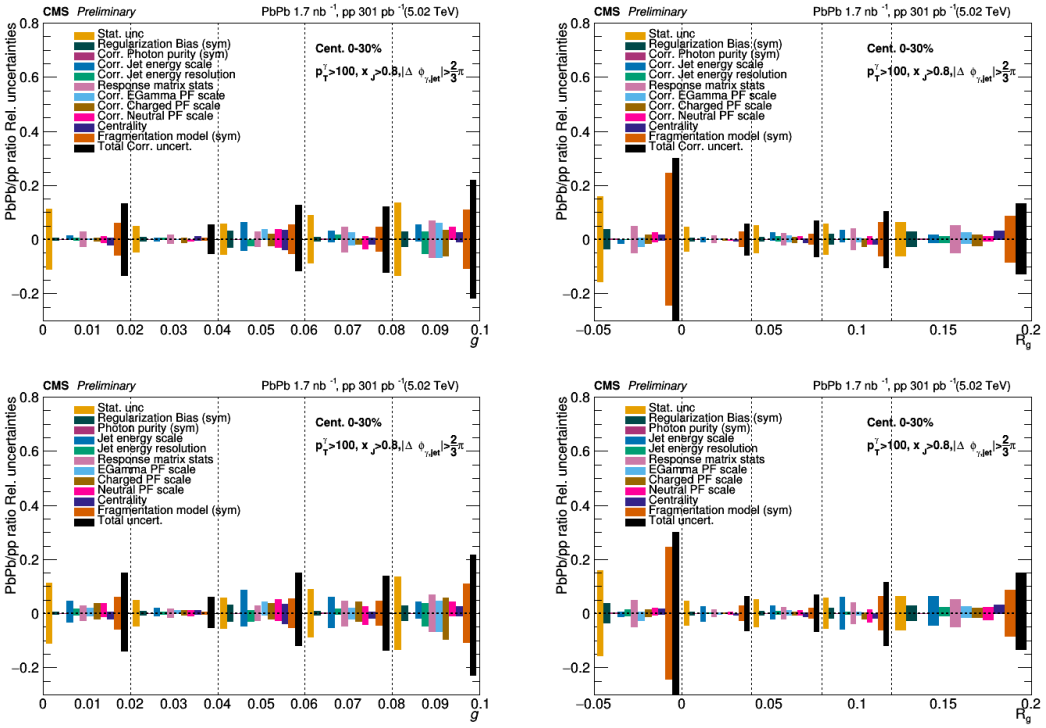


Figure 5.21: Systematic uncertainty breakdown for g (left) and R_g (right) of jets recoiling from photons in the ratio of PbPb to pp collisions for $x_{\gamma j} > 0.8$.A check was performed considering some of the systematic uncertainties as considered fully correlated(upper) and compared to the nominal results where no correlations are considered (lower).

5.10 Feasibility of unregularized unfolding

Unregularized matrix inversion is preferred due to the lack of prior bias. It requires a square response binning (det,true) which is not a restriction in the D'Agostini unfolding method. The diagonality of the response matrix determines the viability of the unregularized matrix inversion and the unfolding method is determined by how ill-conditioned the problem is to guarantee the stability of the unfolding. The degree to which a response matrix K is ill-conditioned can be inferred from its condition number, which is given by $cond(K) = \sigma_{max}/max(0,\sigma_{min})$, where σ_{max} is the largest and σ_{min} is the smallest singular value of K. The condition number represents the degree to which a square response matrix is diagonal, thus smaller condition numbers may represent well-conditioned problems while larger condition numbers tend to represent ill-conditioned problems requiring regularized unfolding. A square response matrix was considered and condition numbers of $\mathcal{O}(> 10)$ were obtained leading to a feasibility study for matrix inversion. We observed oscillations in the solution obtained from matrix inversion as shown in Fig. 5.22.

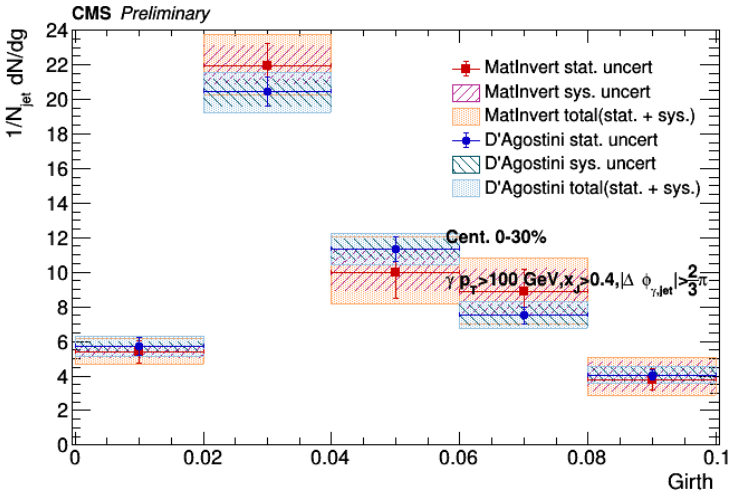


Figure 5.22: Unfolded result for Girth (g) in 0-30% PbPb with D'Agostini and Matrix Inversion.

Comparing the systematic uncertainties between regularized unfolding (D'Agostini) and unregularized unfolding (Matrix Inversion) is shown in Fig. 5.23. We can see that the oscillations get enhanced in the systematic unceratinty leading to larger statistical and systematic uncertainties with matrix inversion.

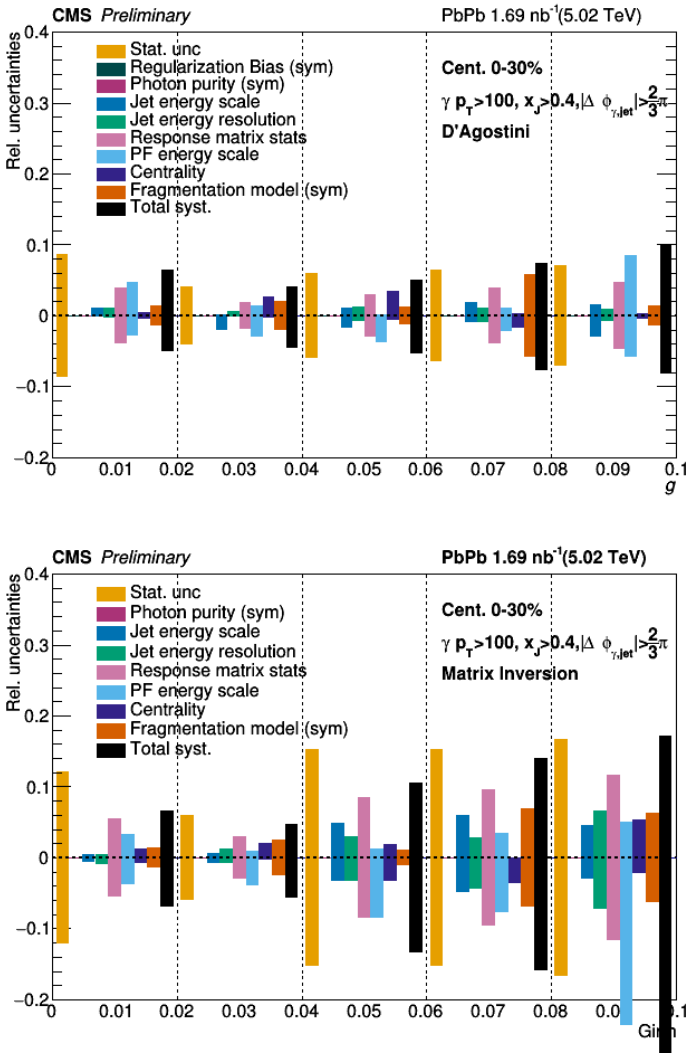


Figure 5.23: Systematic Uncertainty breakdown for Girth (g) unfolded with D'Agostini (left) and Matrix Inversion (right) in 0-30% PbPb

The efficiency and purity for the square binning used for matrix inversion are shown in Table 5.1. We can see a drop in efficiency and purity comparing to Tables 4.2, 4.3, 4.4 and

Table 5.1: Bin-by-bin purity and efficiency for Girth (g) for 0–30% PbPb with $x_{\gamma j} > 0.4$ obtained using MC

Binning	**0–0.02**	**0.02–0.04**	**0.04–0.06**	**0.06–0.08**	**0.08–0.1**
Purity	0.98	0.98	0.95	0.87	0.60
Efficiency	0.98	0.97	0.93	0.90	0.83

4.5 due to restricting the phase space as overflow bins are not included in the unfolding and are corrected with MC based corrections. Figure 5.24 shows the probability matrix for girth in PbPb. The non-diagonalities in the matrix result in larger condition numbers for girth.

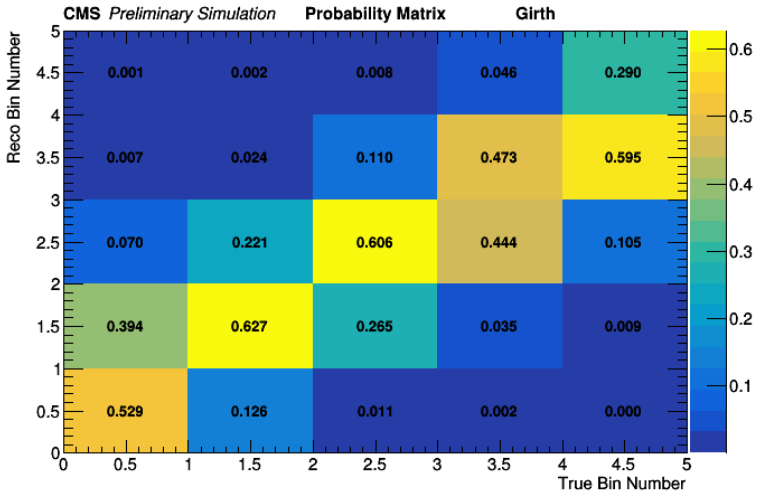


Figure 5.24: Probability Matrix for Girth (g) in 0-30% PbPb with Matrix Inversion.

Thus, matrix inversion is not sufficient to unfold girth to particle level in central PbPb and regularized unfolding is required.

5.11 Unfolding tests

5.11.1 Bottom-line test

A sanity test of the unfolding strategy is the "bottomline test."[1] Unfolding should be such that model discrimination power is not improved at particle-level or at best remains similar as the model discrimination power in smeared space. This is to verify that the unsmearing procedure is not introducing artificial biases in the measured distributions.

χ^2 test

A full bottomline test involves computing the covariance matrix of the data V such that,

$$\chi^2_{\text{smeared}} = (y - K\lambda)^{\mathrm{T}} V_y^{-1} (y - K\lambda) \tag{5.1}$$

$$\chi^2_{\text{unfolded}} = (x - \lambda)^{\mathrm{T}} V_x^{-1} (x - \lambda) \tag{5.2}$$

where y and x are the data before and after unfolding, λ is the model prediction, K is the response matrix and, V_y and V_x are the covariance matrices of the data before and after unfolding. Since the unfolding procedure cannot improve the discrimination power, it is expected that $\chi^2_{\text{unfolded}} < \chi^2_{\text{smeared}}$. In the case of $\chi^2_{\text{unfolded}} \simeq \chi^2_{\text{smeared}}$, it is expected that the unfolding requires little regularization and could be invertible. As shown in the previous section, unregularized unfolding does not yield stable results for the observables in this analysis.

A simplified test was performed considering a unitary covariance matrix to a first approximation as shown in Figs. 5.25 and 5.26 for the observables g and R_{g} for $x_{\gamma\text{j}} > 0.4$ and $x_{\gamma\text{j}} > 0.8$ respectively. As a function of the unfolding iteration, the χ^2_{unfolded} reduces below the χ^2_{smeared} until it plateaus to a constant due to the non-diagonalities of

[1] https://twiki.cern.ch/twiki/bin/viewauth/CMS/ScrecUnfolding#The_bottom_line_test

the response matrix.

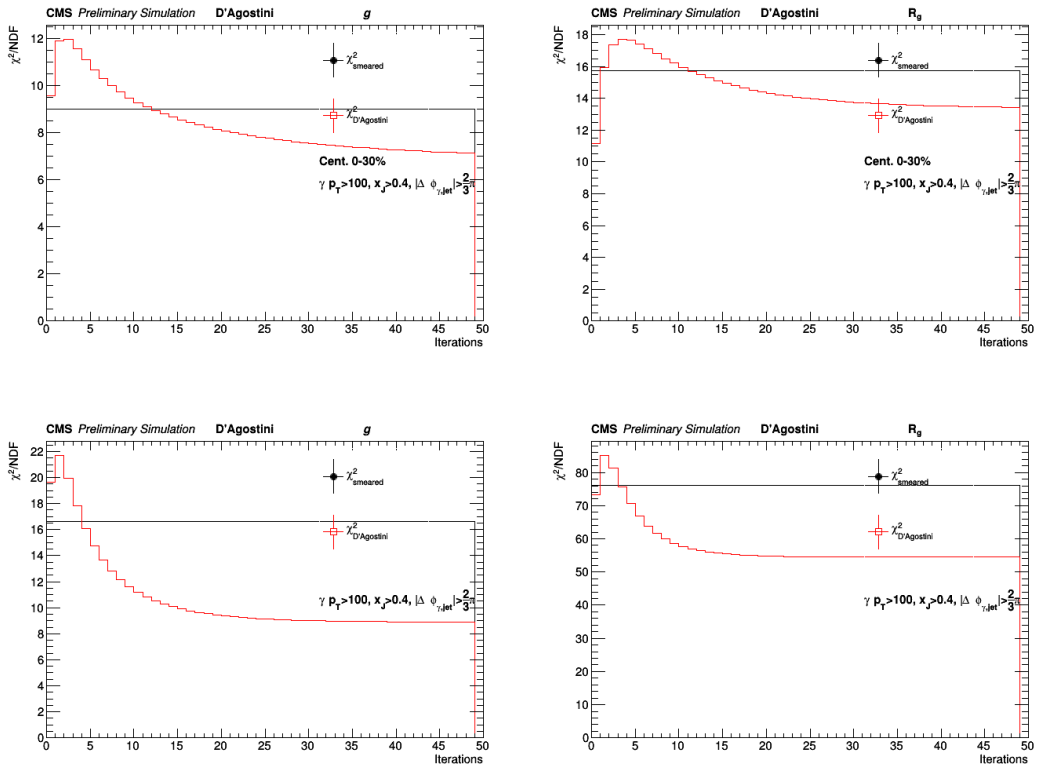


Figure 5.25: Comparison of the χ^2_{unfolded} as a function of iteration with χ^2_{smeared} in a simplified bottomline test for girth g (left) and R_g (right) in central PbPb (upper) and pp (lower) collisions for $x_{\gamma j} > 0.4$. At higher iterations, the $\chi^2_{\text{unfolded}} < \chi^2_{\text{smeared}}$ satisfying the condition

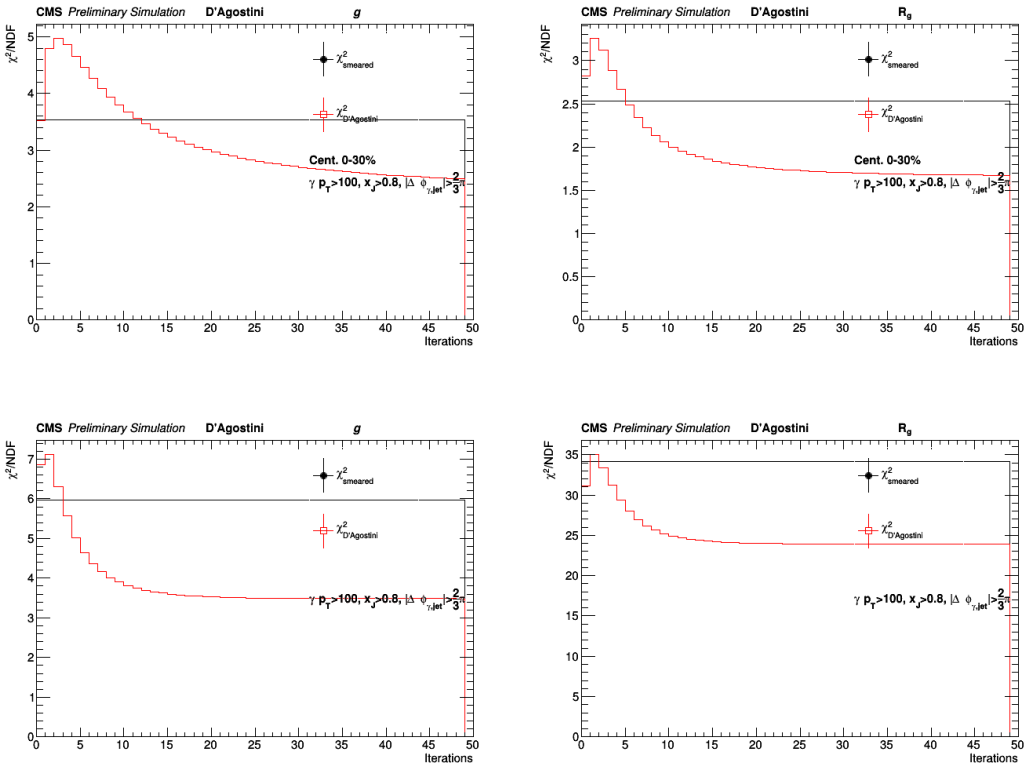


Figure 5.26: Comparison of the χ^2_{unfolded} as a function of iteration with χ^2_{smeared} in a simplified bottomline test for girth g (left) and R_g (right) in central PbPb (upper) and pp (lower) collisions for $x_{\gamma j} > 0.8$. At higher iterations, the $\chi^2_{\text{unfolded}} < \chi^2_{\text{smeared}}$ satisfying the condition

Graphical bottomline test

A complementary graphical bottomline test was performed by comparing the ratio of the MC prediction to data at detector-level and at truth-level, following the procedure in Ref. [182]. We can see the bin-by-bin comparison by taking ratios of the smeared MC to smeared Data and comparing it to true MC to unfolded data. There should not be a significant difference between the 2 ratios as seen in Fig. 5.27 and 5.28. Thus, the bottomline test is considered satisfied.

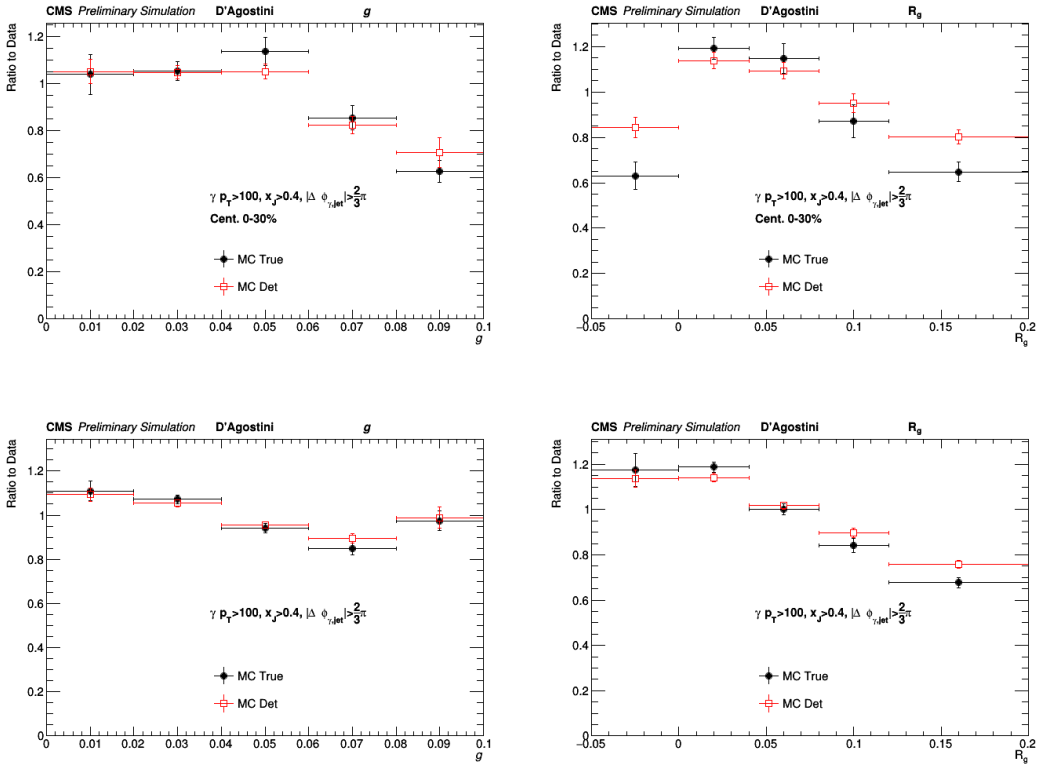


Figure 5.27: Bottom-line test in its graphical version. The ratio of MC to data before (detector-level) and after unfolding (particle-level) is calculated. Unfolding should be such that these two ratios are similar before and after unfolding. The ratios are shown for Girth (g) (left) and R_g(right) for (top to bottom) $0-30\%$ PbPb, and pp. The bottom-line test is satisfied, i.e., the ratio of MC to data has similar trends and sizes before and after unfolding. Only the statistical uncertainties of the input distribution are shown.

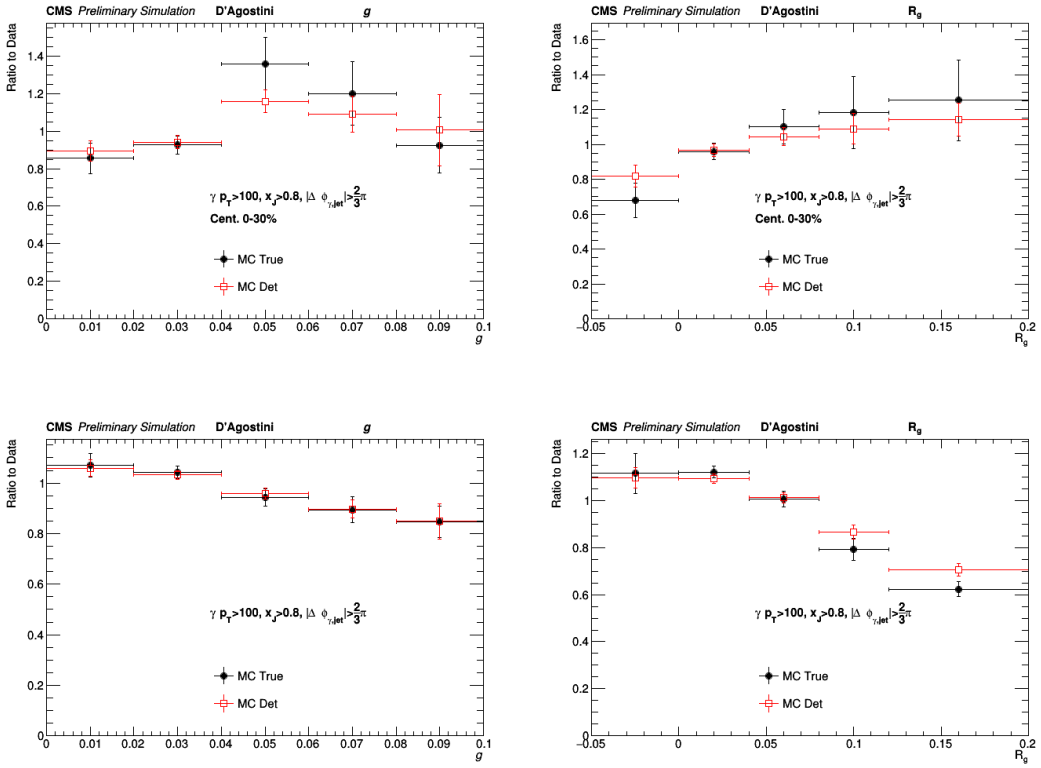


Figure 5.28: Bottom-line test in its graphical version. The ratio of MC to data before (detector-level) and after unfolding (particle-level) is calculated. Unfolding should be such that these two ratios are similar before and after unfolding. The ratios are shown for Girth (g) (left) and R_g(right) for (top to bottom) $0-30\%$ PbPb, and pp. The bottom-line test is satisfied, i.e., the ratio of MC to data has similar trends and sizes before and after unfolding. Only the statistical uncertainties of the input distribution are shown.

5.11.2 Refolding test

The unsmeared distribution is folded back to detector level by matrix multiplication and compared to the input detector-level distribution. This is used to test the hypothesis that the migrations from truth-level to det-level are described in terms of the migration matrix extracted from simulation. These tests are done for both data and MC samples. What we find is that the convergence improves monotonically with each iteration, as expected for the D'Agostini method.

The refolding tests for the data are shown in Figs. 5.29 and 5.30.

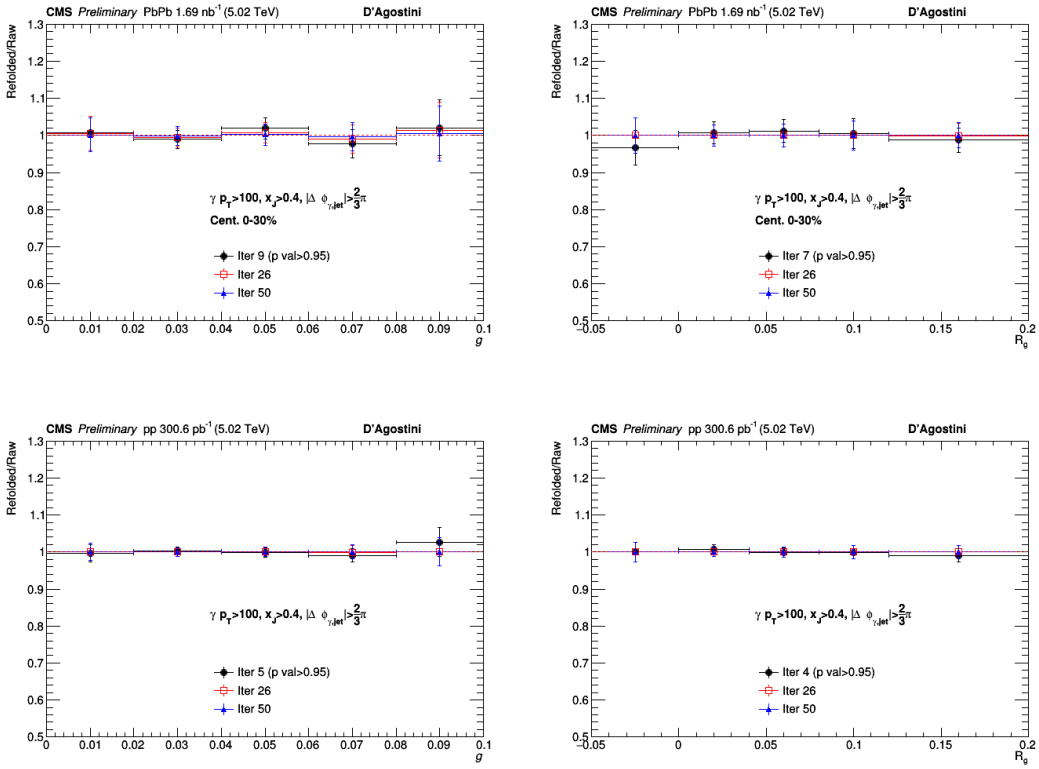


Figure 5.29: Refolding tests for Girth (g) (left) and R_g(right) for (top to bottom) $0-30\%$ PbPb, and pp data unfolded using PYTHIA8 (Embedded) and compared with the input raw distribution at each iteration. We observe closure monotonically with iterations. Only the statistical uncertainties of the input distribution are shown.

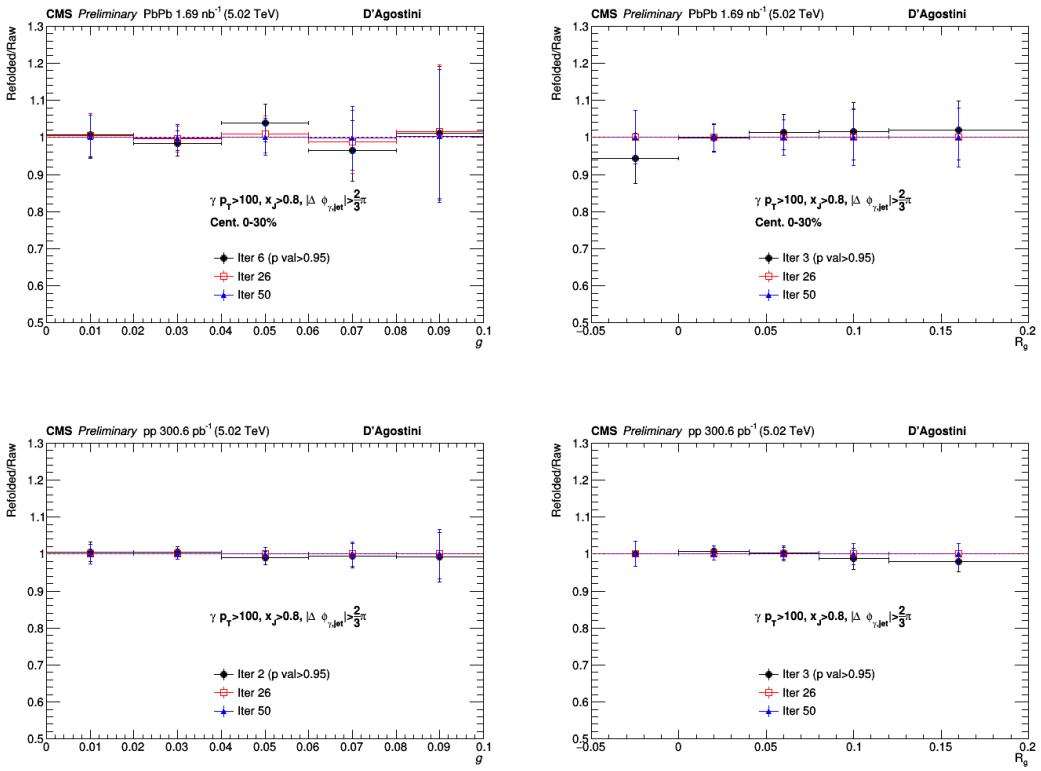


Figure 5.30: Refolding tests for Girth (g) (left) and R_g(right) for (top to bottom) $0-30\%$ PbPb, and pp data unfolded using PYTHIA8 (Embedded) and compared with the input raw distribution at each iteration for $x_J > 0.8$. We observe closure monotonically with iterations. Only the statistical uncertainties of the input distribution are shown.

5.11.3 Trivial test

The unfolded distribution is compared to the truth-level distribution. A trivial test using the full PYTHIA8 Embedded MC sample in both the pseudodata and response to ensure there are no non-closures is performed as shown in Fig. 5.31 and 5.32.

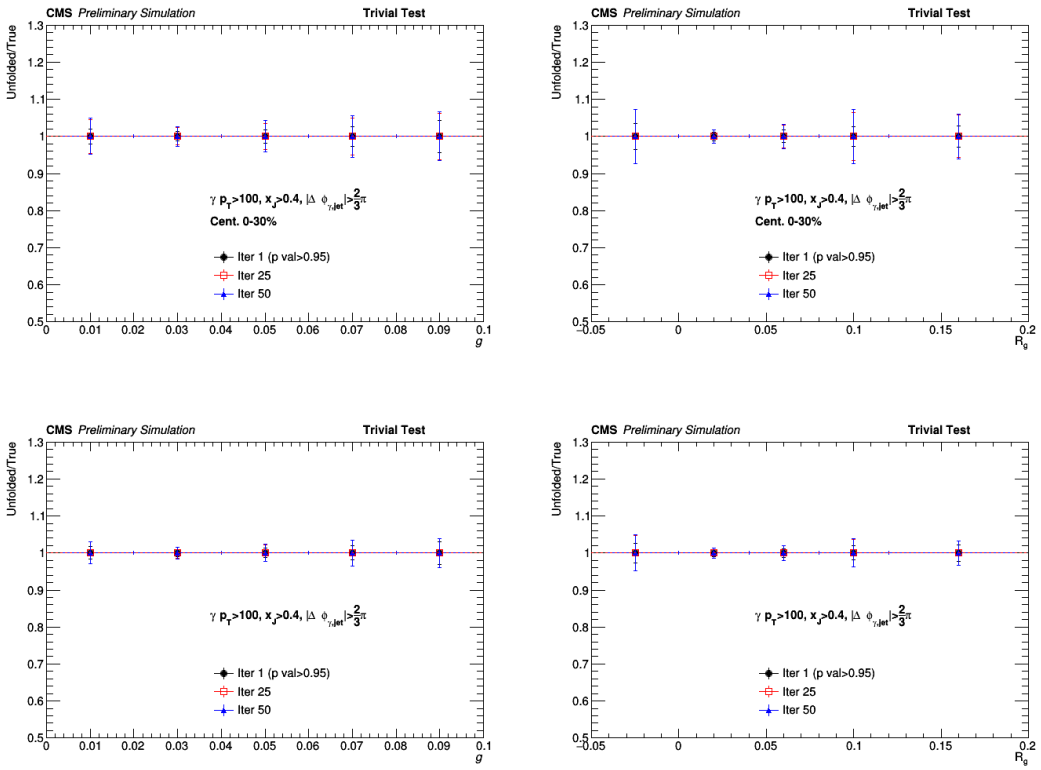


Figure 5.31: Trivial tests in PYTHIA8 (Embedded) MC for Girth (g) (left) and R_g(right) for (top to bottom) $0-30\%$ PbPb, and pp. Since the same samples are used for the pseudodata and the response we expect immediate convergence to 1 and no non closures as we see. Only the statistical uncertainties of the input distribution are shown.

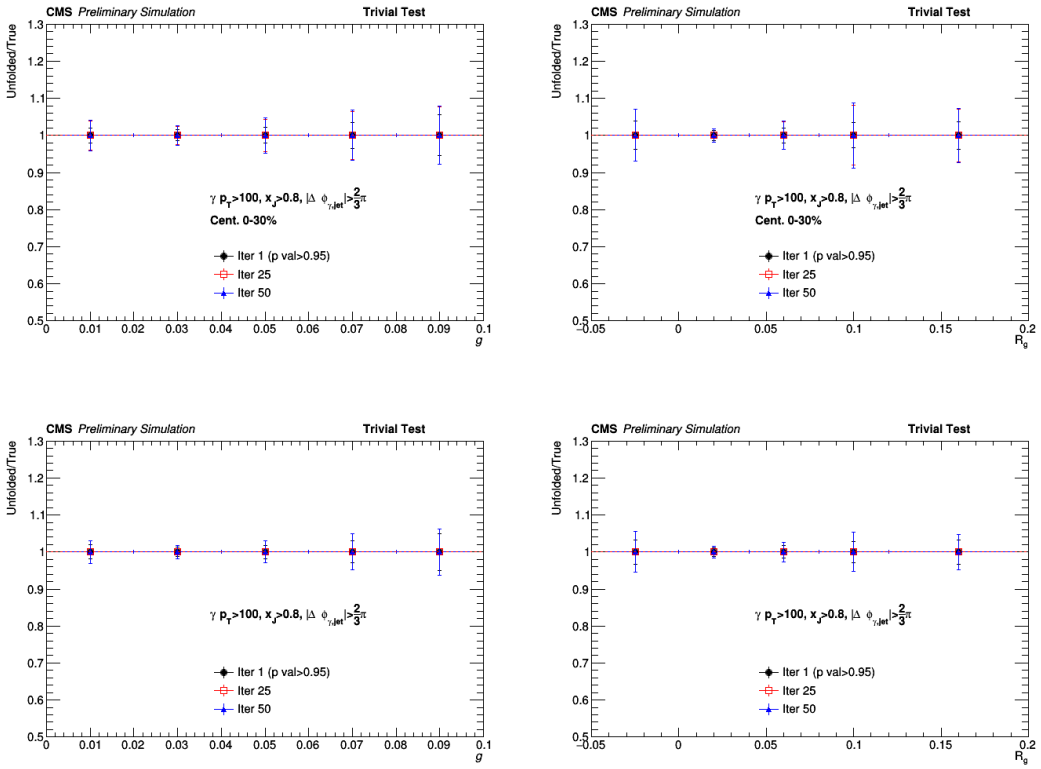


Figure 5.32: Trivial tests in PYTHIA8 (Embedded) MC for Girth (g) (left) and R_g(right) for (top to bottom) $0-30\%$ PbPb, and pp for $x_J > 0.8$. Since the same samples are used for the pseudodata and the response we expect immediate convergence to 1 and no non closures as we see. Only the statistical uncertainties of the input distribution are shown.

5.11.4 Split test

The split tests are performed in a way to ensure that the pseudodata sample is representative of the available statistics in the data. Performing this test, ensures that the data can be unfolded with the given statistical uncertainty if the modeling of the data is perfectly encoded in the response matrix provided. The results are shown in Fig. 5.33 and 5.34.

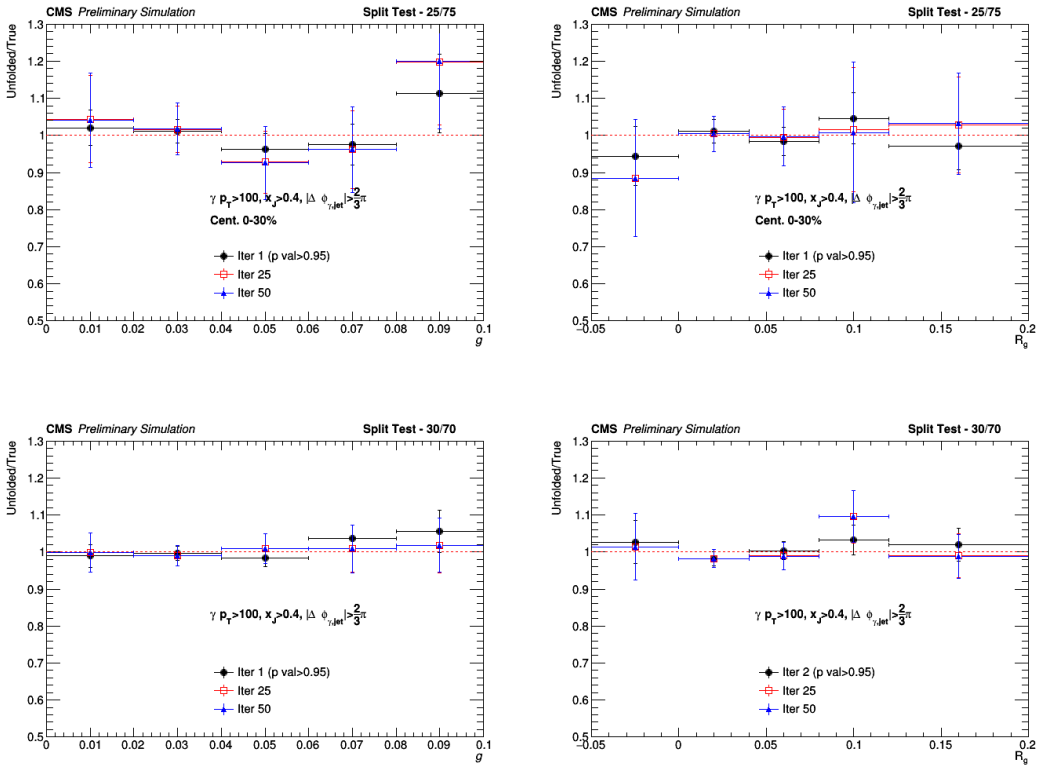


Figure 5.33: Closure tests in split PYTHIA8 (Embedded) MC (statistically independent events for pseudodata and for the corrections) for Girth (g) (left) and R_g(right) for (top to bottom) $0-30\%$ PbPb, and pp for $x_J > 0.4$. It is expected that, since it is the same physics model for both samples, that the truth-level distribution can be recovered within the statistical uncertainties. Only the statistical uncertainties of the input distribution are shown.

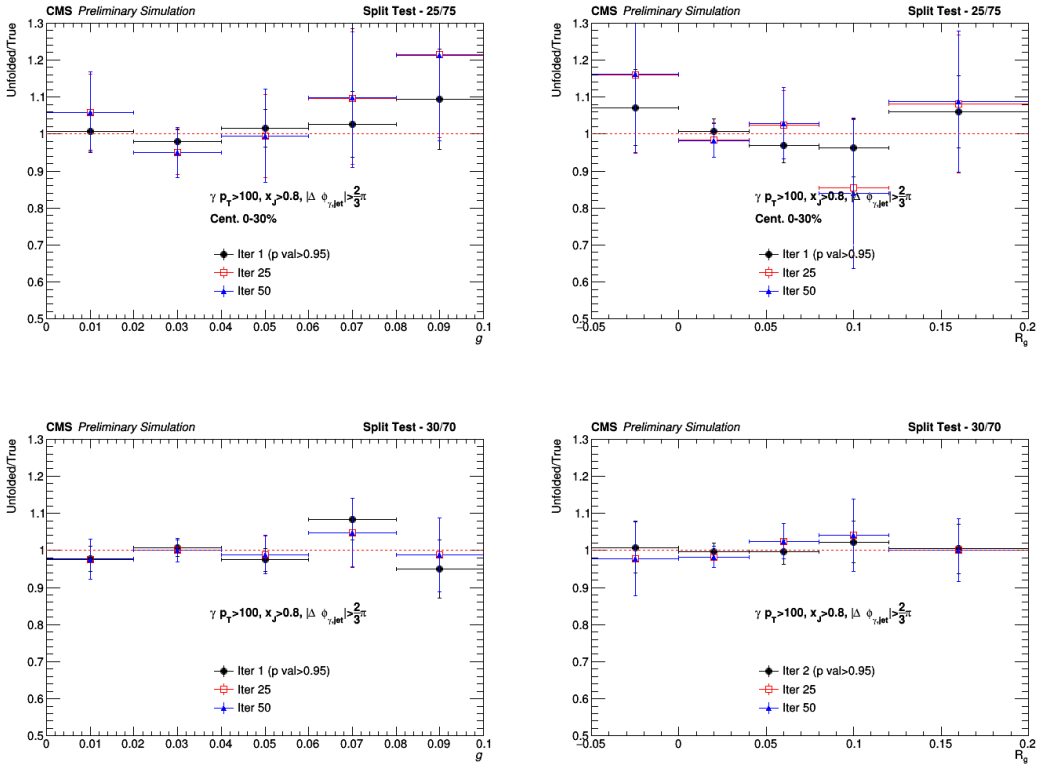


Figure 5.34: Closure tests in split PYTHIA8 (Embedded) MC (statistically independent events for pseudodata and for the corrections) for Girth (g) (left) and R_g(right) for (top to bottom) $0-30\%$ PbPb, and pp for $x_J > 0.8$. It is expected that, since it is the same physics model for both samples, that the truth-level distribution can be recovered within the statistical uncertainties. Only the statistical uncertainties of the input distribution are shown.

5.11.5 Crossfolding test

We consider two further types of closure tests. In one case we split the MC sample and the unfolded solution is compared to the truth. In the other case we use one MC as pseudodata and a different one to fill the response and we check that the unfolded matches the truth. The latter check is done only for pp as we have 2 different parton showers available in PYTHIA8 and HERWIG7 MC in Fig. 5.35 and 5.36. These cross folding tests show the effect of the prior bias due to the dependence of the unfolding on the underlying fragmentation pattern and the non-closures observed are covered by a systematic uncertainty associated to the model dependence.

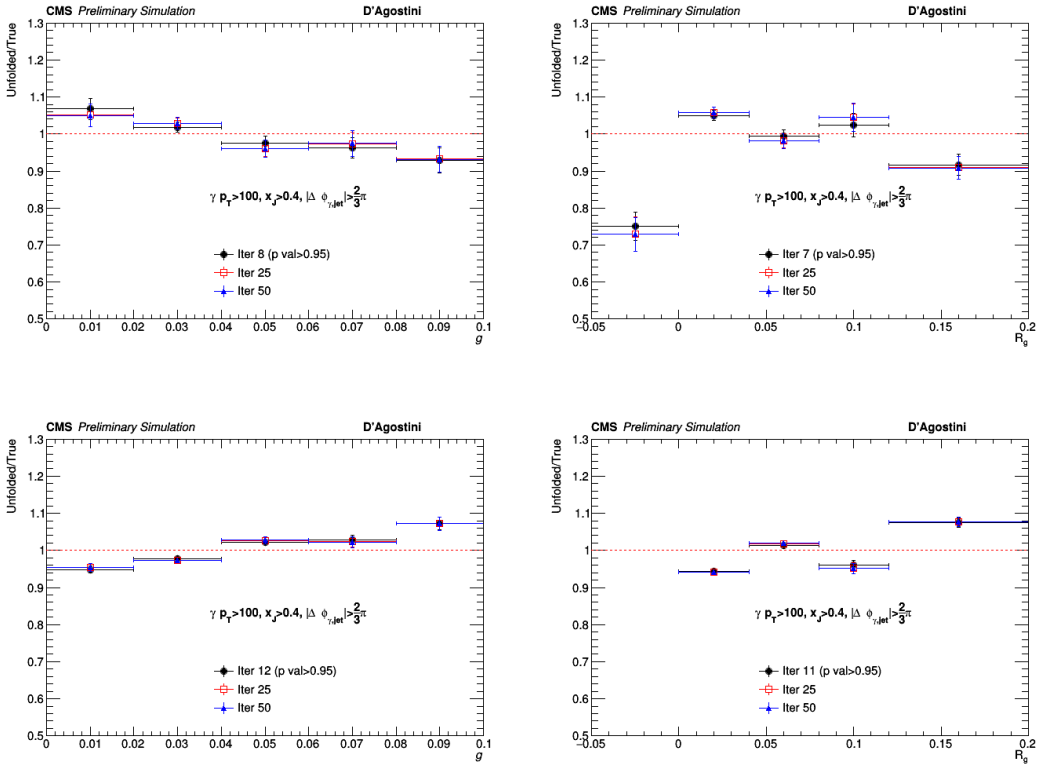


Figure 5.35: Closure tests in pp for Girth (g) (left) and R_g(right) using different MC for the pseudodata and for the response. Upper panel: The PYTHIA8 distribution is unfolded with HERWIG7 and compared with the input true-level distribution at each iteration. Lower panel: Same test is done, interchanging PYTHIA8 and HERWIG7. Only the statistical uncertainties of the input distribution are shown. Non-closures are observed, which are due to the existence of a prior bias in the corrections. These non-closures are covered with the systematic uncertainties associated to the model dependence. We can also see that the non-closure trends are opposite when interchanging the raw and input true-level distributions.

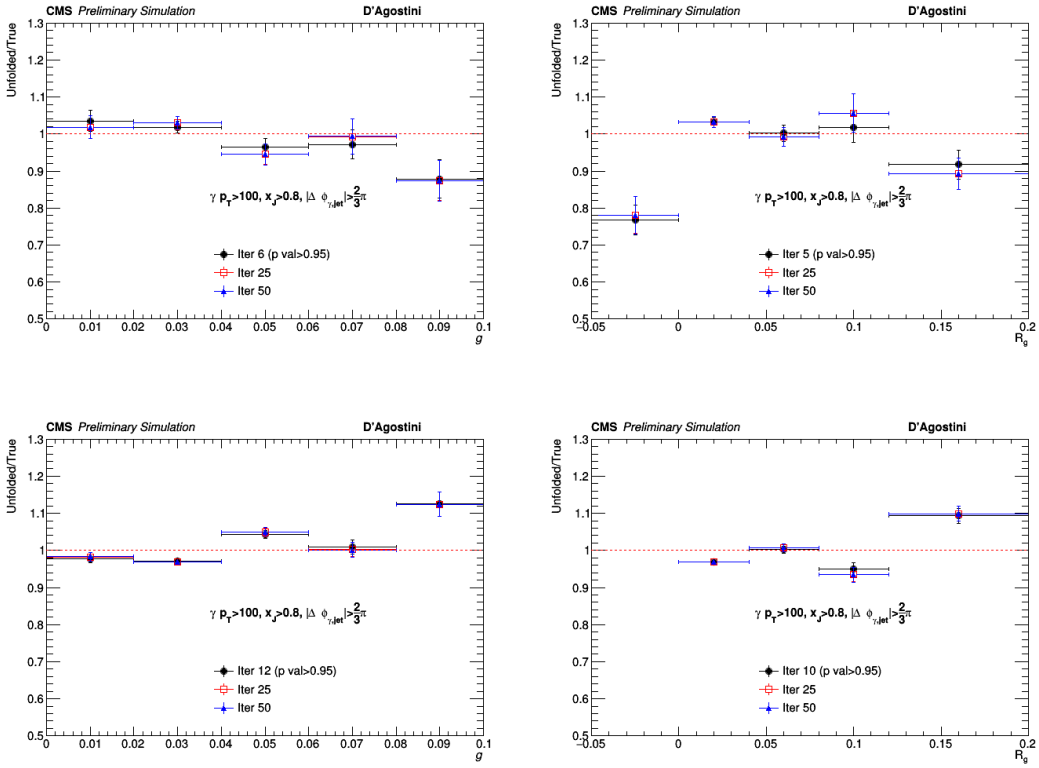


Figure 5.36: Closure tests in pp for Girth (g) (left) and R_g(right) using different MC for the pseudodata and for the response. upper panel: The PYTHIA8 distribution is unfolded with HERWIG7 and compared with the input true-level distribution at each iteration. lower panel: Same test is done, interchanging PYTHIA8 and HERWIG7. Only the statistical uncertainties of the input distribution are shown. Non-closures are observed, which are due to the existence of a prior bias in the corrections. These non-closures are covered with the systematic uncertainties associated to the model dependence. We can also see that the non-closure trends are opposite when interchanging the raw and input true-level distributions.

5.12 Narrowing in pp

Previous measurements of narrowing in inclusive jet substructure measurements as discussed in Section 1.3.4 could be caused due to the inherent ambiguity in the jet clustering procedure. Jet algorithms utilize a resolution parameter i.e. the jet radius R, that defines the opening angle of the jet to provide control over the divergences of QCD in perturbation theory. Thus, energy loss can be a expected result in measurements of jets due to out of cone radiation for these limited size jet radii [172]. Quarks and gluons, due to their different radiation pattern also have a different degree of out of cone radiation which can lead to differing degrees of energy loss as shown in Fig. 1.29. For a given 100 GeV parton, a p_{T} shift can be estimated at leading order, given the parton is a quark or gluon as follows:

$$\langle \delta p_{\mathrm{T}} \rangle_q = -C_F \frac{\alpha_S p_{\mathrm{T}}}{\pi} \ln\left(\frac{1}{R}\right)\left(2\ln 2 - \frac{3}{8}\right) \tag{5.3}$$

$$\implies \langle \delta p_{\mathrm{T}} \rangle_q (100) \approx \text{6-8 GeV} \tag{5.4}$$

$$\langle \delta p_{\mathrm{T}} \rangle_g = -\frac{\alpha_S p_{\mathrm{T}}}{\pi} \ln\left(\frac{1}{R}\right)\left[C_A(2\ln 2 - \frac{43}{96}) + T_R n_f \frac{7}{48}\right] \tag{5.5}$$

$$\implies \langle \delta p_{\mathrm{T}} \rangle_g (100) \approx \text{18-20 GeV} \tag{5.6}$$

The color factors for quarks ($C_F = 4/3$) and gluons ($C_A = 3$) affect the p_{T} shift leading to larger shifts for gluons than quarks. There is an inverse dependence on the radius of the jet as well for the energy loss. This difference can be checked in photon tagged jet substructure in pp collisions by selecting jets with harder fragmentation ($x_{\gamma \mathrm{j}} > 0.8$) and comparing with a looser selection ($x_{\gamma \mathrm{j}} > 0.4$). Figure 5.37 shows the comparison of unfolded data reproducing narrowing as discussed in Section 1.3.4 within a pp collision system without the QGP medium effects. The p_{T} shift effect on R_{g} for quarks and

gluons can be separately studied in simulated events as shown in Fig. 5.38 with a larger shift seen in gluons than quarks.

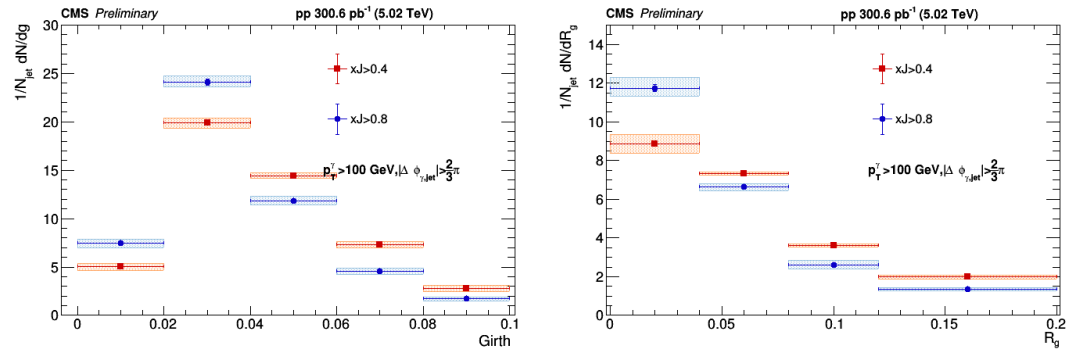


Figure 5.37: Narrowing observed for Girth (g) (left) and R_g(right) in pp data unfolded using PYTHIA8 between $x_{\gamma j} > 0.4$ and $x_{\gamma j} > 0.8$

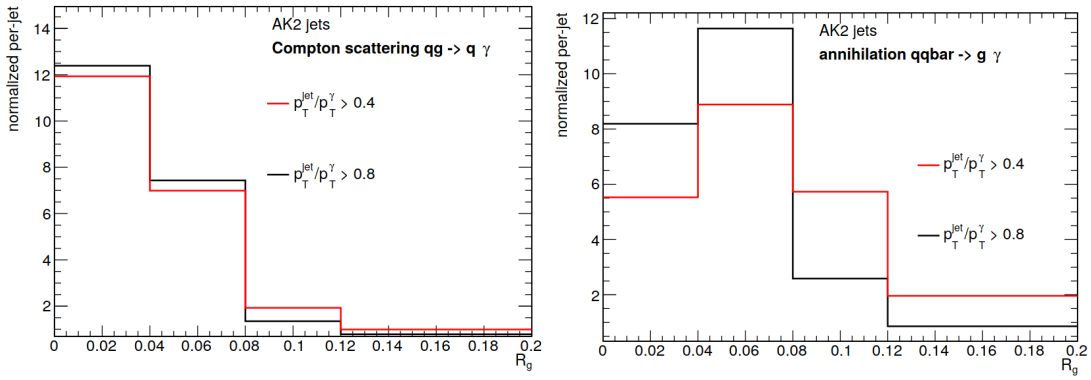


Figure 5.38: Simulation of energy shifts in R_g for quarks (left) and gluons (right) by comparing between $x_{\gamma j} > 0.4$ and $x_{\gamma j} > 0.8$. Gluons show a larger shift as expected compared to quarks

Chapter 6

Results

The unfolded distributions of g and R_g of photon-tagged jets in pp collisions are shown in Figs. 6.1 and 6.2 for $x_{\gamma j} > 0.4$ and $x_{\gamma j} > 0.8$, respectively. In pp collisions, one can use the $x_{\gamma j} > 0.4$ and 0.8 categories to better understand out-of-cone radiation effects, which lead to an average shift in jet p_T [172]. On average, gluon-initiated jets experience a larger p_T shift from out-of-cone radiation than quark-initiated jets [172], so selecting events based on $x_{\gamma j} > 0.4$ and 0.8 leads to different parton flavor compositions.

The pp measurement is compared with MC simulation results from PYTHIA 8.230 CP5, PYTHIA 8.303 with the VINCIA shower [183], PYTHIA 8.303 with the DIRE shower [184], HERWIG 7.2.2 with an angular-ordered shower using tune CH3, and HERWIG 7.2.0 with a dipole shower [179, 180]. The DIRE shower implements a p_T-ordered color dipole shower, where the radiator–spectator particle pairs evolve simultaneously, and it includes higher-order corrections, such as triple-collinear or double-soft parton emissions. The version of VINCIA in PYTHIA 8.303 uses the antenna sector shower formalism [183, 185, 186]. The branching kernels, known as antenna functions, treat coherent sums of parton pairs without requiring a separation into radiators and spectators. Both VINCIA and DIRE have their respective set of tuning parameters in PYTHIA8. The dipole parton shower available in HERWIG7 uses the Catani–Seymour dipole factorization formalism [187]. We also compare the "hybrid" model predictions for vacuum radiation with

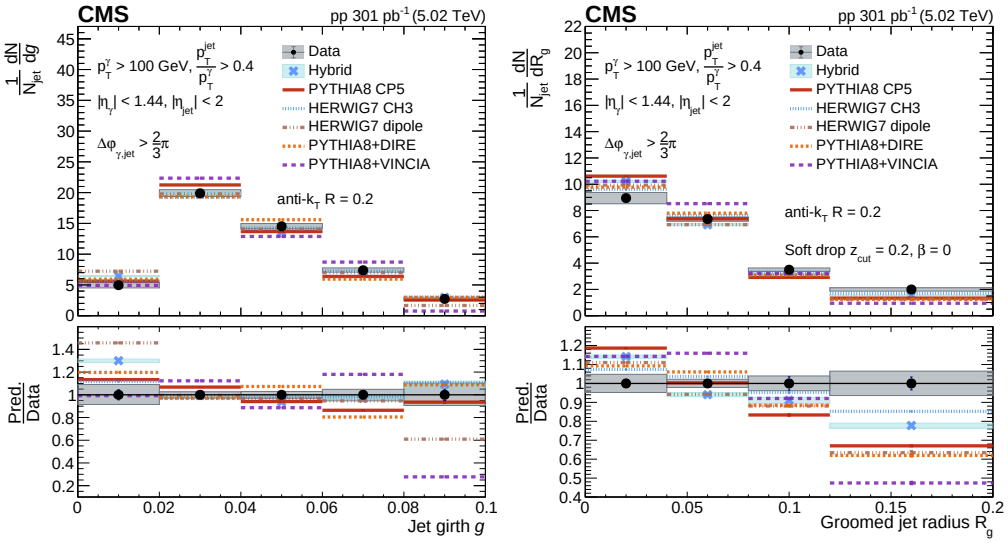


Figure 6.1: Unfolded distributions of jet girth g (left) and groomed jet radius R_g (right) of photon-tagged jets in pp collisions for $x_{\gamma j} \equiv p_T^{jet}/p_T^{\gamma} > 0.4$. The upper panels show the comparison of the observable in pp collisions and predictions from simulated events. The lower panels show the corresponding ratios of the MC calculations and data. The bands represent the total uncertainties, whereas the vertical bars represent the statistical uncertainties.

the data [53], which consists of PYTHIA8 generator predictions with the Monash tune as a baseline [188], but without multiple parton interactions, a slightly larger p_T cutoff for the final-state radiation, and with the PDF set NNPDF2.3 QCD+QED LO [171]. All the predictions presented here use matrix elements at LO in perturbation theory for the hard scattering.

Differences between pp unfolded distributions and the MC simulated predictions are of the order of 20% in the bulk of the distributions. The largest discrepancies are at small and large values of g and at large values of R_g. This indicates that the jets measured in data are, on average, broader than the jets in simulation. The PYTHIA8 CP5 predictions describe the g distribution for both $x_{\gamma j} > 0.4$ and 0.8 selections within the uncertainties, but not R_g. The hybrid model predictions for pp have similar trends as PYTHIA8 CP5. The best global description is provided by the HERWIG7 CH3 generator for both $x_{\gamma j} > 0.4$ and 0.8 categories in the g and R_g variables, whereas the HERWIG7

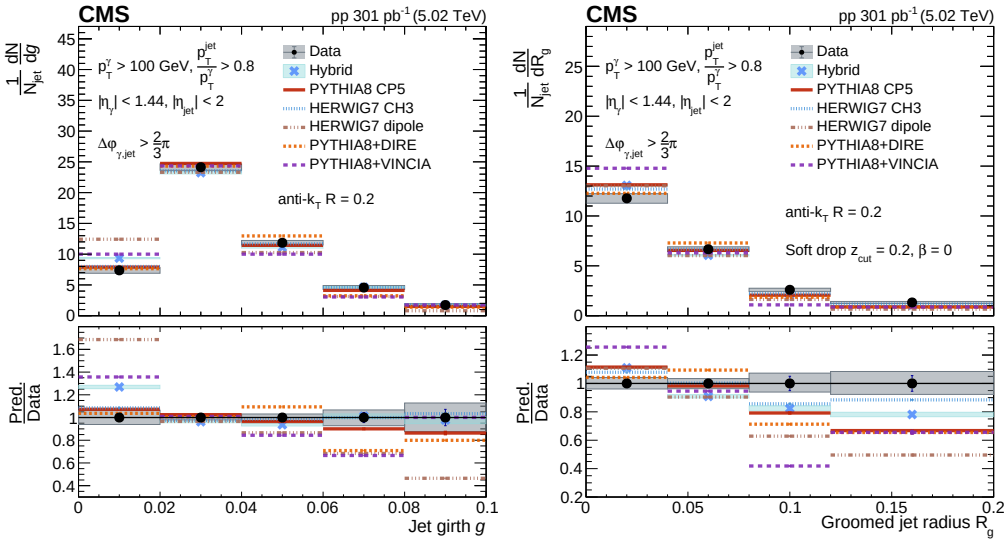


Figure 6.2: Unfolded distributions of jet girth g (left) and groomed jet radius R_g (right) of photon-tagged jets in pp collisions for $x_{\gamma j} \equiv p_T^{jet}/p_T^{\gamma} > 0.8$. The upper panels show the comparison of the observable in pp collisions and predictions from simulated events. The lower panels show the corresponding ratios of the MC calculations and data. The bands represent the total experimental uncertainties, whereas the vertical bars represent the statistical uncertainties.

generator with a dipole shower tends to have the largest differences with the data. Recent measurements of jet substructure in Z+jet and dijet events in pp collisions [21] suggest that the jet substructure modeling depends on details of the quark and gluon jet composition of the sample. Photon+jet events are expected to have a larger quark jet fraction relative to inclusive jets for a similar jet p_T domain. Thus, the present measurement can help constrain the modeling of quark jet fragmentation in simulations, which can be useful to improve the modeling of the vacuum shower baseline used in predictions for medium-induced jet modification.

The unfolded distributions of g and R_g for photon-tagged jets in PbPb collisions compared with those of pp collisions are shown in Figs. 6.3 and 6.4 for $x_{\gamma j} > 0.4$ and 0.8. The ratios of the unfolded distributions of PbPb to pp are shown in Fig. 6.5 for $x_{\gamma j} > 0.4$ and in Fig. 6.6 for $x_{\gamma j} > 0.8$. The ratios of PbPb to pp distributions are used to identify potential medium-induced modifications of the jet shower. For photon-tagged jets

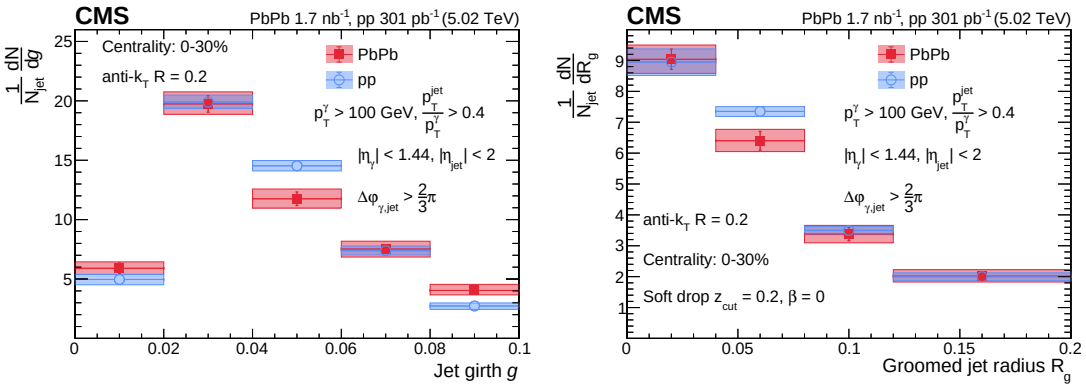


Figure 6.3: Unfolded distributions of jet girth g (left) and groomed jet radius R_g (right) of photon-tagged jets in PbPb and pp collisions for $x_{\gamma j} \equiv p_T^{jet}/p_T^\gamma > 0.4$ (selecting both more and less quenched jets). The bands represent the total uncertainties, whereas the vertical bars represent the statistical uncertainties.

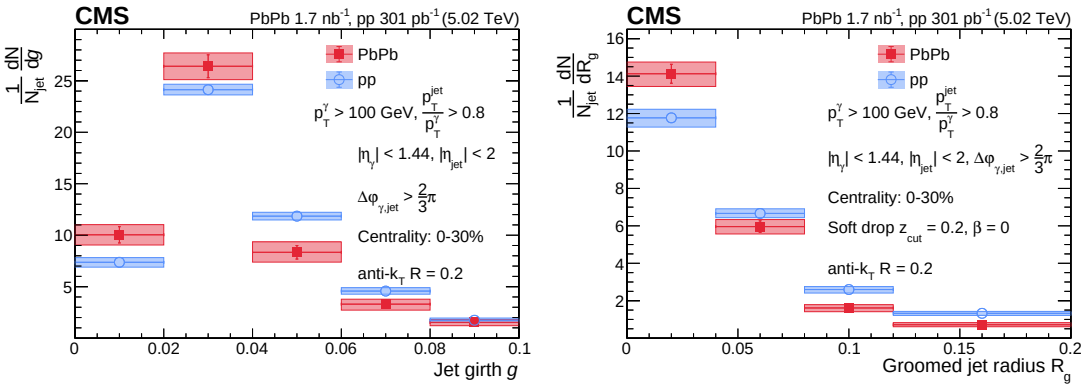


Figure 6.4: Unfolded distributions of jet girth g (left) and groomed jet radius R_g (right) of photon-tagged jets in PbPb and pp collisions for $x_{\gamma j} \equiv p_T^{jet}/p_T^\gamma > 0.8$ (selecting less quenched jets). The bands represent the total uncertainties, whereas the vertical bars represent the statistical uncertainties.

with $x_{\gamma j} > 0.4$ in Fig. 6.3, which corresponds to a more inclusive jet selection where both quenched and less quenched jets are selected, we observe no narrowing of the angular structure of jets produced in PbPb collisions relative to the distributions in pp collisions within the experimental uncertainties. For g, there are hints of a broadening of the substructure of the jet in PbPb collisions at large values of g, whereas R_g is consistent with pp within the experimental uncertainties. On the other hand, by selecting less quenched jets via $x_{\gamma j} > 0.8$ in Fig. 6.4, which is a less inclusive jet selection, we observe a narrowing of the angular structure of jets in PbPb collisions when compared with pp collisions, contrary to the trend observed in events selected with $x_{\gamma j} > 0.4$. Figures. 6.5 and 6.6 compare the measurements of the ratio of the unfolded distributions of PbPb to pp with the hybrid model [53]. The model is a "hybrid" approach of weak and strong coupling approximations to describe medium-induced jet modifications. It includes the modeling of energy loss due to the strong coupling between the partons and the medium, large-angle deflections of partons transversing the QGP, referred to as "Moliére elastic scatterings" [69, 73], and the nonperturbative back reaction of the medium [74], also known as medium response. These effects are denoted as "elastic" and "wake" in Figs. 6.5 and 6.6, and the absence of such effects as "no elastic" or "no wake." In the model, the medium resolution length is controlled by the parameter L_{res} [189], which corresponds to the minimum transverse length between two color-connected partons for the medium to resolve them separately. Three values of L_{res} are considered: $L_{\text{res}} = 0, 2/(\pi T)$, and ∞, where T is the temperature of the medium in the model. The $L_{\text{res}} = 0$ value corresponds to the incoherence limit where all the radiators are resolved by the medium. The $L_{\text{res}} = 2/(\pi T)$ corresponds to the expectation that L_{res} should be of the same order of magnitude as the Debye screening length [189] and it represents an intermediate scenario where only a fraction of the radiators interact with the medium. The $L_{\text{res}} = \infty$ value corresponds to the full coherence limit, where radiators are not resolved individually and the jet interacts with the medium as a sin-

gle color charge. In the case of elastic scatterings with the medium, the present model accounts for them for $L_{\mathrm{res}} = 0$.

Some conclusions can be drawn from the comparison of these theory predictions from the hybrid model to the experimental data. First, the contribution of the wake is negligible, as shown in Figs. 6.5 and 6.6, because of the small R parameter used in the measurement. Second, there is no consistent choice of parameters that can describe both the $x_{\gamma\mathrm{j}} > 0.4$ and 0.8 categories simultaneously. For $x_{\gamma\mathrm{j}} > 0.4$ in Fig. 6.5, the incoherence case $L_{\mathrm{res}} = 0$ overestimates the narrowing, while the model describes the data better when the Molière elastic scatterings are added, which broadens the jet, or when the limiting case of $L_{\mathrm{res}} = \infty$ is considered. The $x_{\gamma\mathrm{j}} > 0.4$ data sample alone cannot be used to separate the contribution of color coherence and elastic scatterings, but when used together with the unfolded distributions of $x_{\gamma\mathrm{j}} > 0.8$, one can better separate the consequences of such effects, as shown in Fig. 6.6. The incoherence limit $L_{\mathrm{res}} = 0$ is favored in the $x_{\gamma\mathrm{j}} > 0.8$ category, and both the effects of elastic scatterings and the effect of nonzero L_{res} values result in a slight overestimation of the normalized yield at large R_{g} values. For g in Figs. 6.5 and 6.6, the results of the model comparison to the data are qualitatively the same as for R_{g}, with an underestimation of the PbPb-to-pp ratio at large values of g from all setups for $x_{\gamma\mathrm{j}} > 0.4$. Since the categories $x_{\gamma\mathrm{j}} > 0.4$ and 0.8 enable the selection of different degrees of jet energy loss, the present measurement has sensitivity on how the different physical mechanisms affect the radiation pattern of the jet as well as the jet energy loss. Thus, the present measurement can be used to optimize the model parameter choice together with existing measurements that are sensitive to other aspects of medium-induced jet modifications.

The measurement using photon-tagged jets is complementary to previous measurements in inclusive jet events. In inclusive jet measurements, the comparison between jets in PbPb and pp collisions is done at the same reconstructed jet p_{T}, but due to the

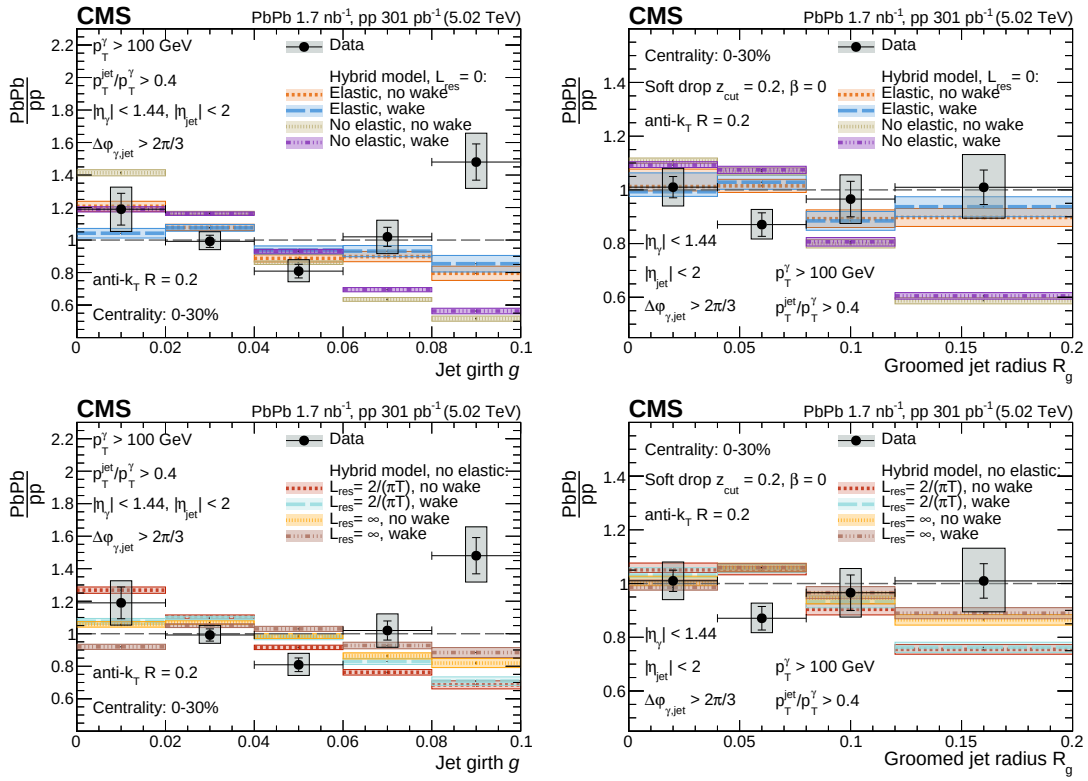


Figure 6.5: Ratio of the normalized yields of PbPb to pp data for jet girth g (left) and groomed jet radius R_g (right) of photon-tagged jets in PbPb and pp collisions for $x_{\gamma j} \equiv p_T^{jet}/p_T^\gamma > 0.4$ (selecting both more and less quenched jets). The data are compared with the hybrid model predictions for $L_{res} = 0$ (upper) and for nonzero values of L_{res} without elastic scattering (lower). The bands around the data points represent the total experimental uncertainties, whereas the vertical bars represent the statistical uncertainties. The uncertainties in the PbPb-to-pp ratio have been obtained assuming the PbPb and pp measurements are uncorrelated. The bands around the theory predictions represent the statistical uncertainties of the prediction.

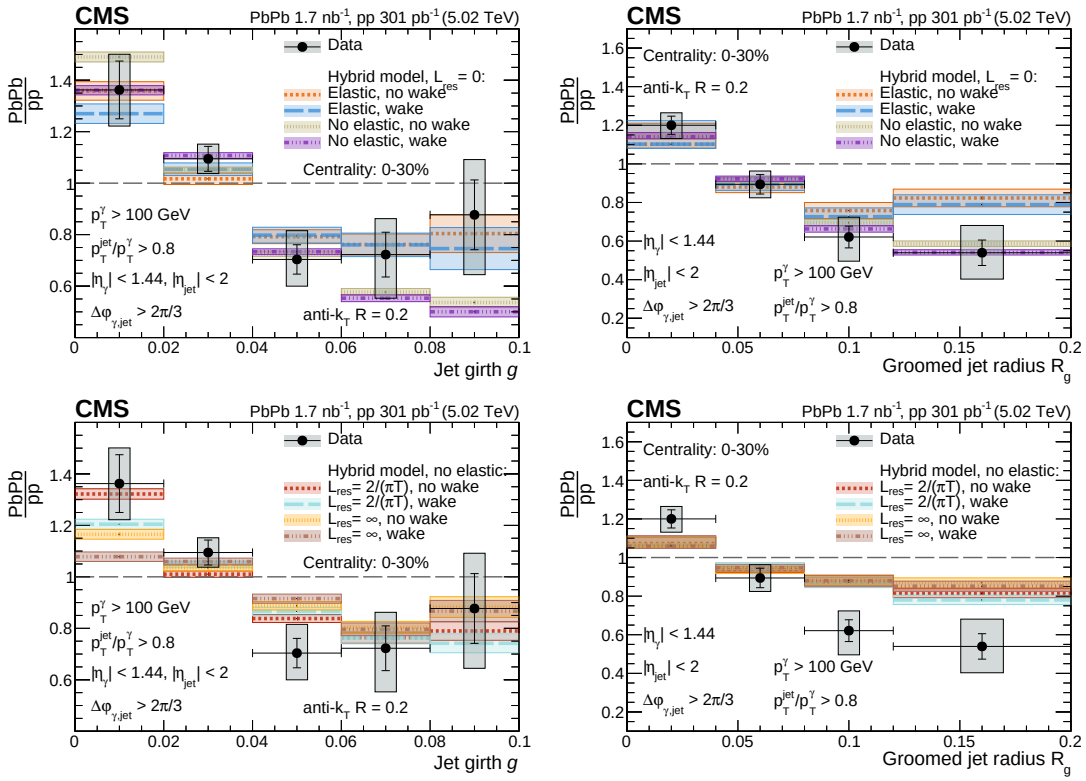


Figure 6.6: Ratio of the normalized yields of PbPb to pp for jet girth g (left) and groomed jet radius R_g (right) of photon-tagged jets in PbPb and pp collisions for $x_{\gamma j} \equiv p_T^{jet}/p_T^\gamma > 0.8$ (selecting less quenched jets). The data are compared with the hybrid predictions for $L_{res} = 0$ (upper) and nonzero values of L_{res} without elastic scatterings (lower). The bands around the data points represent the total experimental uncertainties, whereas the vertical bars represent the statistical uncertainties. The uncertainties in the PbPb-to-pp ratio have been obtained assuming the PbPb and pp measurements are uncorrelated. The bands around the theory predictions represent the statistical uncertainties of the prediction.

jet energy loss that occurs in PbPb collisions, the comparison is not done for the same initial parton energies. In contrast to the trends observed by the ALICE and ATLAS Collaborations for R_{g} or g in inclusive jet events [81, 96], we do not observe a narrowing of the substructure of jets in R_{g} or g within the experimental uncertainties when selecting jets with $x_{\gamma\mathrm{j}} > 0.4$ and $p_{\mathrm{T}}^{\gamma} > 100\,\mathrm{GeV}$. One possible explanation is that, after reducing the selection bias effect due to jet energy loss, the sample of events being analyzed in photon-tagged events consists not only of narrow unquenched jets but also of broader, and thus more quenched, jets. By selecting less quenched jets via $x_{\gamma\mathrm{j}} > 0.8$, we observe an effective narrowing in the angular substructure of jets produced in PbPb collisions, qualitatively similar to the distributions measured by the ALICE and ATLAS Collaborations [81, 96]. We verified that the conclusions of the study are the same even after imposing a requirement on $x_{\gamma\mathrm{j}} < 1$ to reduce the contribution of bremsstrahlung photons.

6.1 Summary

In summary, we report the first measurements girth g and the groomed jet radius R_{g} of jets recoiling against isolated photons in lead-lead (PbPb) and proton-proton (pp) collisions. The analysis uses PbPb and pp collision data, both at a nucleon-nucleon center-of-mass energy of 5.02 TeV. The distributions are unfolded to the particle level in order to facilitate comparisons between experiments and with theoretical predictions.

The transverse momentum p_{T} of isolated photons (p_{T}^{γ}) can be used as a proxy for the p_{T} of the high-virtuality parton that initiates the shower of the recoiling jet. This enables the disentanglement of the potential modification of the momentum and angular substructure of jets due to the interactions with the medium from the selection bias effects that can originate from jet energy loss. This is done using the transverse momentum imbalance, defined as the ratio of the hardest recoil jet p_{T} ($p_{\mathrm{T}}^{\mathrm{jet}}$) and p_{T}^{γ}, $x_{\gamma\mathrm{j}} \equiv p_{\mathrm{T}}^{\mathrm{jet}} / p_{\mathrm{T}}^{\gamma}$.

It is found that jets with $p_{\mathrm{T}}^{\mathrm{jet}}/p_{\mathrm{T}}^{\gamma} > 0.8$, i.e., those that closely balance the photon p_{T}^{γ}, are narrower in PbPb than in pp collisions. Relaxing the selection to include jets with $p_{\mathrm{T}}^{\mathrm{jet}}/p_{\mathrm{T}}^{\gamma} > 0.4$ reduces the narrowing of the angular structure of jets in PbPb relative to the pp reference. These observations suggest that selection bias effects play an important role in the interpretation of the modification of the angular scales of jets in terms of medium-induced effects. The measured distributions are compared with calculations based on a hybrid strong and weak coupling model to describe medium-induced jet modifications. According to model predictions, the R_{g} and g distributions are not very sensitive to medium response effects or to variations of the medium resolution length. However, changes in the modeling of Molière elastic scatterings have an effect of 10–40% at large values of g and R_{g}. This shows the ability of the data to constrain the impact of Molière scatterings in a way that is effectively factorized from the effects of the wake and the medium resolution length.

Medium-induced jet modifications are commonly assessed by comparing jets and their substructure at the same reconstructed p_{T} in PbPb and pp collisions, which in the former case corresponds to the momentum of the jet after its interactions with the quark gluon plasma. These interactions are expected to broaden the jet and reduce its energy. Thus, in an inclusive jet measurement, when comparing populations of jets in PbPb and pp within the same measured jet p_{T} window, a selection bias can lead to an effective narrowing of the angular structure of jets in PbPb relative to pp. One possibility is that the population of jets that were initially broader (hence, more strongly quenched jets) has migrated to lower jet energies, whereas the population of narrower jets (less strongly quenched jets) remains. Thus, events with high-p_{T} jets recoiling against energetic isolated photons can be used to better constrain genuine medium modifications of the jet shower, complementing measurements in inclusive jet production.

Chapter 7

Résumé en francais

Le Modèle Standard de la physique des particules est le cadre le plus réussi pour décrire les particules élémentaires et leurs interactions. Une des théories fondamentales constituant le Modèle Standard est connue sous le nom de Chromodynamique Quantique (QCD), qui décrit les interactions entre les quarks et les gluons. Ceux-ci possèdent une propriété appelée charge de couleur, par laquelle ils participent à l'interaction forte. Pour comprendre l'interaction forte, au cours des dernières décennies, nous avons étudié la diffusion d'un faisceau de particules par rapport à une cible fixe ou à un autre faisceau de particules. L'énergie de l'interaction détermine les échelles de distance qui peuvent être résolues, ce qui conduit à l'étude de collisions à plus haute énergie ces dernières années.

Une manière d'étudier la QCD à des échelles de distances courtes est l'étude de la production de jets, des jets de particules hautement collimés. Les jets sont définis par l'algorithme utilisé pour leur reconstruction et le schéma d'énergie. Dans les collisions hadroniques, les jets reconstruits avec l'algorithme anti-k_{T} sont couramment utilisés en raison de leur résilience aux effets de bruit de fond. Les jets ont été étudiés comme des objets monolithiques dans le passé, mais des avancées récentes dans les modèles théoriques et les développements expérimentaux ont permis d'explorer l'interaction forte à travers l'évolution du jet. Ceci est étudié à l'aide de différents observables de

sous-structure de jets.

Des mesures des observables de sous-structure de jets, telles que la girth g et le groomed des jets (R_g), sont effectuées en résonance avec des photons isolés dans des collisions plomb-plomb (PbPb) et proton-proton (pp) au LHC. L'analyse utilise des échantillons de données PbPb et pp collectés avec le détecteur CMS en 2017 et 2018, tous deux à une énergie de centre de masse nucléon-nucléon de 5.02 TeV, avec des luminosités intégrées de 301 pb^{-1} et 1.7 nb^{-1}, respectivement. Ces mesures fournissent un aperçu des propriétés du plasma de quarks et de gluons (QGP), un milieu chaud et dense créé dans les collisions PbPb. Les modifications des jets dans le milieu QGP sont étudiées par rapport à une référence fournie par les événements de collisions pp. Les observables g et R_g offrent une mesure quantitative de la structure plus ou moins étroit des jets. Les mesures de ces observables dans les événements de jets inclusifs ont montré une sous-structure de jets plus étroit dans les collisions d'ions lourds comparées aux collisions pp pour les jets reconstruits avec la même impulsion transverse du jet p_T^{jet} dans les mesures précédentes. Une mesure de ces observables dans les événements de jets tagués par des photons offre une interprétation complémentaire de la perte d'énergie des jets dans le milieu QGP en considérant l'impulsion transverse du photon p_T^{γ} comme un proxy pour le p_T de la particule initiant la cascade de jets.

Les événements doivent contenir un photon avec une impulsion transverse $p_T^{\gamma} > 100\,\text{GeV}$ et au moins un jet opposé en azimut par rapport au photon et avec une impulsion transverse p_T^{jet} telle que $p_T^{jet}/p_T^{\gamma} > 0.4$. Les distributions mesurées de R_g et g sont dépliées au niveau des particules, ce qui facilite la comparaison entre les résultats PbPb et pp ainsi qu'avec les prédictions théoriques. Il est constaté que les jets avec $p_T^{jet}/p_T^{\gamma} > 0.8$, c'est-à-dire ceux qui équilibrent étroitement le p_T^{γ} du photon, sont plus étroits dans PbPb que dans pp. En relâchant la sélection pour inclure les jets avec $p_T^{jet}/p_T^{\gamma} > 0.4$, le rétrécissement de la structure angulaire des jets dans PbPb par rapport à la référence

pp est réduit. Cela montre que les biais de sélection associés à la perte d'énergie des jets jouent un rôle important dans l'interprétation des mesures de sous-structure des jets.

Cette thèse est structurée pour fournir tous les détails nécessaires dans le contexte de ces études. Au Chapitre 1, une discussion sur les mesures actuelles et la motivation physique de cette analyse est présentée. Le Chapitre 2 fournit des détails sur le détecteur CMS utilisé pour la collecte des données. Des informations détaillées sur la reconstruction des objets physiques tels que les photons et les jets utilisés dans cette analyse à partir des informations du détecteur sont présentées au Chapitre 3. Les Chapitres 4 et 5 discutent de la procédure d'analyse suivie et des vérifications effectuées à chaque étape pour garantir des résultats scientifiques robustes, respectivement. Les résultats de l'analyse effectuée dans cette thèse sont présentés au Chapitre 6 [95].

Appendix A

Technical details (CMS internal use)

The complete code used for this photon tagged jet substructure analysis presented in this thesis is available in this github repository: `https://github.com/hbharadwaj/HeavyIonAnalysis`.

Tables A.1, A.2 and A.3 lists the datasets used in this analysis, intended for internal use by CMS members.

Table A.1: Data samples

Sample	Dataset path
PbPb 2018 (5.02 TeV)	`/HIHardProbes/HIRun2018A-04Apr2019-v1/AOD`
pp 2017 (5.02 TeV)	`/HighEGJet/Run2017G-17Nov2017-v2/AOD`

The global tags `103X_dataRun2_Prompt_v2` and `94X_dataRun2_ReReco_EOY17_v6` are used for PbPb and pp data respectively. For MC, we use `103X_upgrade2018_realistic_HI_v11` and `94X_mc2017_realistic_forppRef5TeV` for PbPb and pp. The CMSSW release for PbPb and pp are CMSSW_10_3_3_patch1 and CMSSW_9_4_10. The centrality tag for PbPb data is

`CentralityTable_HFtowers200_DataPbPb_periHYDJETshape_run2v1033p1x01_offline`

Our configurations are based on the HIForest setup. The conditions consumed by the HiForest (JEC, centrality) are not pulled from GlobalTag, but rather are overwritten. The jet energy resolution in MC is further smeared using the hybrid method with

Table A.2: PbPb MC samples

Dataset path
/QCDPhoton_pThat-XXX_TuneCP5_HydjetDrumMB_5p02TeV_Pythia8/ HINPbPbAutumn18DR-mva98_103X_upgrade2018_realistic_HI_v11-v1/AODSIM XXX = 30,50,80
/QCDPhoton_pThat-XXX_TuneCP5_HydjetDrumMB_5p02TeV_Pythia8/ HINPbPbAutumn18DR-mva98_103X_upgrade2018_realistic_HI_v11-v2/AODSIM XXX = 120,170
/QCDPhoton_pThat-XXX_Filter30GeV_TuneCP5_HydjetDrumMB_5p02TeV_Pythia8/ HINPbPbAutumn18DR-mva98_103X_upgrade2018_realistic_HI_v11-v1/AODSIM XXX = 30,50,80,120
/QCDPhoton_pThat-XXX_Filter30GeV_TuneCP5_HydjetDrumMB_5p02TeV_Pythia8/ HINPbPbAutumn18DR-mva98_103X_upgrade2018_realistic_HI_v11-v2/AODSIM XXX = 170

Table A.3: pp MC samples

Dataset path
/QCDPhoton_pThat-XXX_TuneCP5_5p02TeV_pythia8/ RunIIpp5Spring18DR-94X_mc2017_realistic_forppRef5TeV_v1-v1/AODSIM XXX = 30,50,80,120,170
/QCDPhoton_pThat-XXX_Filter30GeV_TuneCP5_5p02TeV_pythia8/ RunIIpp5Spring18DR-94X_mc2017_realistic_forppRef5TeV_v1-v2/AODSIM XXX = 30
/QCDPhoton_pThat-XXX_Filter30GeV_TuneCP5_5p02TeV_pythia8/ RunIIpp5Spring18DR-94X_mc2017_realistic_forppRef5TeV_v1-v1/AODSIM XXX = 50,80,120,170
/QCD_PthatGT15_TuneCH3_5p02TeV_herwig7/ RunIIpp5Spring18DR-94X_mc2017_realistic_forppRef5TeV_v2-v2/AODSIM

AK2PF collections:

- **2018 PbPb MC:** Autumn18_RunD_V7b_MC_PtResolution_AK4PF
- **2017 pp MC:** Fall17_V3b_MC_PtResolution_AK4PF

We use the approved luminosity sections selected for each year and collision mode:

- **2018 PbPb:** Cert_326381-327564_HI_PromptReco_Collisions18_JSON.txt
- **2017 pp:** Cert_306546-306826_5TeV_EOY2017ReReco_Collisions17_JSON.txt

Appendix B

List of Abbreviations

SM	Standard Model
QED	Quantum Electrodynamics
QCD	Quantum Chromodynamics
CMS	Compact Muon Solenoid
LHC	Large Hadron Collider
pQCD	perturbative QCD
FF	Fragmentation Function
PDF	Parton Distribution Function
WTA	Winner-Take-All
UE	Underlying Event
IRC	Infrared and Collinear safe
IC-SM	Iterative Cone with Split-Merge procedure
IC-PR	Iterative Cone with Progressive Removal
SISCone	Seedless Infrared Safe Cone
CA	Cambridge-Aachen
SD	Soft Drop
QGP	Quark-Gluon Plasma
MC	Monte Carlo

AdS/CFT	anti-de Sitter/Conformal Field Theory
PS	Proton Synchroton
ISR	Intersecting Storage Rings
SPS	Super Proton Synchrotron
LHC	Large Hadron Collider
RF	Radio Frequency
PSB	Proton Synchrotron Booster
LIER	Low Energy Ion Ring
ALICE	A Large Ion Collider Experiment
RHIC	Relativistic Heavy Ion Collider
ATLAS	A Toroidal LHC ApparatuS
LHCb	LHC beauty
LHCf	LHC forward
TOTEM	TOTal cross-section, Elastic scattering and diffraction dissociation Measurement at the LHC
MOEDAL	MOnopole and Exotics Detector at the LHC
pp	proton-proton
PbPb	lead-lead
TIB	Tracker Inner Barrel
TOB	Tracker Outer Barrel
TID	Tracker Inner Disk
TEC	Tracker End Cap
ECAL	Electromagnetic Calorimeter
EB	ECAL Barrel
EE	ECAL Endcap
ES	ECAL Preshower disks
APD	Avalanche Photo Diodes

VPT	Vacuum Photo Triodes
HCAL	Hadronic Calorimeter
HB	HCAL barrel
HE	HCAL Endcap
HO	HCAL Outer barrel
HF	HCAL Forward
HPD	Hybrid Photo Diode
SiPM	Silicon Photo Multipliers
DT	Drift Tubes
CSC	Cathode Strip Chambers
RPC	Resistive Plate Chambers
MB	Minimum Bias
L1	Level 1 trigger
HLT	High-Level Trigger
FPGA	Field-Programmable Gate Array
TP	Trigger Primitives
GT	Global Trigger
PF	Particle Flow
GSF	Gaussian-sum filter
GED	Global Event Description
SC	Supercluster
QGP	Quark-Gluon Plasma
DSCB	Double Sided Crystal Ball
CS	Constituent Subtraction
JES or JEC	Jet energy scale corrections
JER	Jet Energy Resolution
MPF	Momentum Projection Fraction

PD	Primary datasets
LO	Leading Order
NLO	next to Leading Order
PYQUEN	PYthia QUENched
JETPHOX	JET-PHOton/hadron X sections
HERWIG	Hadron Emission Reactions With Interfering Gluons
SHERPA	Simulation of High-Energy Reactions of Particles
GEANT	GEometry ANd Tracking
VINCIA	VIrtual Numerical Collider with Interleaved Antennae
HYDJET	HYDrodynamics plus JETs

Figures

Tables

Bibliography

[1] Claude Amsler et al. "Review of Particle Physics". In: *Phys. Lett. B* 667 (2008), pp. 1–1340. DOI: 10.1016/j.physletb.2008.07.018 (cit. on p. 7).

[2] Gregg Jaeger. "The Elementary Particles of Quantum Fields". In: *Entropy* 23.11 (2021), p. 1416. DOI: 10.3390/e23111416 (cit. on p. 8).

[3] *File:Standard Model of Elementary Particles.svg - Wikipedia — en.wikipedia.org*. https://en.wikipedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg. [Accessed 03-06-2024] (cit. on p. 8).

[4] *Volume of an Atom and Nucleus - Nuclear Power — nuclear-power.com*. https://www.nuclear-power.com/nuclear-power/reactor-physics/atomic-nuclear-physics/atomic-nuclear-structure/volume-atom-nucleus/. [Accessed 18-06-2024] (cit. on p. 12).

[5] G. Dissertori. "The Determination of the Strong Coupling Constant". In: *Adv. Ser. Direct. High Energy Phys.* 26 (2016), pp. 113–128. DOI: 10.1142/9789814733519_0006. arXiv: 1506.05407 [hep-ex] (cit. on p. 13).

[6] A. Adare et al. "Inclusive double-helicity asymmetries in neutral-pion and eta-meson production in $\vec{p}+\vec{p}$ collisions at $\sqrt{s} = 200$ GeV". In: *Phys. Rev. D* 90.1 (2014), p. 012007. DOI: 10.1103/PhysRevD.90.012007. arXiv: 1402.6296 [hep-ex] (cit. on p. 14).

[7] John C. Collins, Davison E. Soper, and George F. Sterman. "Factorization for Short Distance Hadron - Hadron Scattering". In: *Nucl. Phys. B* 261 (1985), pp. 104–142. DOI: 10.1016/0550-3213(85)90565-6 (cit. on p. 15).

[8] John C. Collins, Davison E. Soper, and George F. Sterman. "Soft Gluons and Factorization". In: *Nucl. Phys. B* 308 (1988), pp. 833–856. DOI: 10.1016/0550-3213(88)90130-7 (cit. on p. 15).

[9] Matt Dobbs and Jorgen Beck Hansen. "The HepMC C++ Monte Carlo event record for High Energy Physics". In: *Comput. Phys. Commun.* 134 (2001), pp. 41–46. DOI: 10.1016/S0010-4655(00)00189-2 (cit. on p. 16).

[10] Gavin Salam. *PhD-level lectures — gsalam.web.cern.ch*. https://gsalam.web.cern.ch/teaching/PhD-courses.html. [Accessed 24-06-2024] (cit. on pp. 16, 17, 19).

[11] Yuri L. Dokshitzer et al. *Basics of perturbative QCD*. 1991 (cit. on p. 18).

[12] Bo Andersson et al. "Parton fragmentation and string dynamics". In: *Phys. Rept.* 97 (1983), p. 31. DOI: 10.1016/0370-1573(83)90080-7 (cit. on pp. 18, 111).

[13] B. R. Webber. "A QCD Model for Jet Fragmentation Including Soft Gluon Interference". In: *Nucl. Phys. B* 238 (1984), p. 492. DOI: 10.1016/0550-3213(84)90333-X (cit. on pp. 19, 111).

[14] R. Keith Ellis, W. James Stirling, and B. R. Webber. *QCD and collider physics*. Vol. 8. Cambridge University Press, Feb. 2011. ISBN: 978-0-511-82328-2, 978-0-521-54589-1. DOI: 10.1017/CBO9780511628788 (cit. on p. 21).

[15] Gerald C. Blazey et al. "Run II jet physics". In: *Physics at Run II: QCD and Weak Boson Physics Workshop: Final General Meeting*. May 2000, pp. 47–77. arXiv: hep-ex/0005012 (cit. on p. 21).

[16] G. Arnison et al. "Hadronic Jet Production at the CERN Proton - anti-Proton Collider". In: *Phys. Lett. B* 132 (1983), p. 214. DOI: 10.1016/0370-2693(83)90254-X (cit. on p. 21).

[17] Gavin P. Salam and Gregory Soyez. "A Practical Seedless Infrared-Safe Cone jet algorithm". In: *JHEP* 05 (2007), p. 086. DOI: 10.1088/1126-6708/2007/05/086. arXiv: 0704.0292 [hep-ph] (cit. on p. 21).

[18] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. "The anti-k_T jet clustering algorithm". In: *JHEP* 04 (2008), p. 063. DOI: 10.1088/1126-6708/2008/04/063. arXiv: 0802.1189 [hep-ex] (cit. on pp. 21, 22, 101, 119).

[19] Yuri L. Dokshitzer et al. "Better jet clustering algorithms". In: *JHEP* 08 (1997), p. 001. DOI: 10.1088/1126-6708/1997/08/001. arXiv: hep-ph/9707323 [hep-ph] (cit. on pp. 21, 119).

[20] Stan Bentvelsen and Irmtraud Meyer. "The Cambridge jet algorithm: Features and applications". In: *Eur. Phys. J. C* 4 (1998), pp. 623–629. DOI: 10.1007/s100520050232. arXiv: hep-ph/9803322 (cit. on p. 21).

[21] Armen Tumasyan et al. "Study of quark and gluon jet substructure in Z +jet and dijet events from pp collisions". In: *JHEP* 01 (2022), p. 188. DOI: 10.1007/JHEP01(2022)188. arXiv: 2109.03340 [hep-ex] (cit. on pp. 24, 134, 154, 181).

[22] Simone Marzani, Gregory Soyez, and Michael Spannowsky. "Looking inside jets: an introduction to jet substructure and boosted-object phenomenology". In: *Lect. Notes Phys.* 958 (2019), p. 1. DOI: 10.1007/978-3-030-15709-8. arXiv: 1901.10342 [hep-ph] (cit. on pp. 24, 45).

[23] Hsi-Ming Chang et al. "Calculating Track Thrust with Track Functions". In: *Phys. Rev. D* 88 (2013), p. 034030. DOI: 10.1103/PhysRevD.88.034030. arXiv: 1306.6630 [hep-ph] (cit. on p. 24).

[24] Hsi-Ming Chang et al. "Calculating Track-Based Observables for the LHC". In: *Phys. Rev. Lett.* 111 (2013), p. 102002. DOI: 10.1103/PhysRevLett.111.102002. arXiv: 1303.6637 [hep-ph] (cit. on p. 24).

[25] Benjamin T. Elder and Jesse Thaler. "Aspects of Track-Assisted Mass". In: *JHEP* 03 (2019), p. 104. DOI: 10.1007/JHEP03(2019)104. arXiv: 1805.11109 [hep-ph] (cit. on p. 24).

[26] Jesse Thaler and Ken Van Tilburg. "Identifying Boosted Objects with N-subjettiness". In: *JHEP* 03 (2011), p. 015. DOI: 10.1007/JHEP03(2011)015. arXiv: 1011.2268 [hep-ph] (cit. on p. 24).

[27] Andrew J. Larkoski, Gavin P. Salam, and Jesse Thaler. "Energy Correlation Functions for Jet Substructure". In: *JHEP* 06 (2013), p. 108. DOI: 10.1007/JHEP06(2013)108. arXiv: 1305.0007 [hep-ph] (cit. on p. 24).

[28] Ian Moult, Lina Necib, and Jesse Thaler. "New Angles on Energy Correlation Functions". In: *JHEP* 12 (2016), p. 153. DOI: 10.1007/JHEP12(2016)153. arXiv: 1609.07483 [hep-ph] (cit. on p. 24).

[29] Andrew J. Larkoski, Jesse Thaler, and Wouter J. Waalewijn. "Gaining (Mutual) Information about Quark/Gluon Discrimination". In: *JHEP* 11 (2014), p. 129. DOI: 10.1007/JHEP11(2014)129. arXiv: 1408.3122 [hep-ph] (cit. on p. 24).

[30] Aram Hayrapetyan et al. "Measurement of the primary Lund jet plane density in proton-proton collisions at $\sqrt{s}$ = 13 TeV". In: *JHEP* 05 (2024), p. 116. DOI: 10.1007/JHEP05(2024)116. arXiv: 2312.16343 [hep-ex] (cit. on pp. 25, 26).

[31] André H. Hoang et al. "Nonperturbative Corrections to Soft Drop Jet Mass". In: *JHEP* 12 (2019), p. 002. DOI: 10.1007/JHEP12(2019)002. arXiv: 1906.11843 [hep-ph] (cit. on p. 27).

[32] Jonathan M. Butterworth et al. "Jet substructure as a new Higgs search channel at the LHC". In: *Phys. Rev. Lett.* 100 (2008), p. 242001. DOI: 10.1103/PhysRevLett.100.242001. arXiv: 0802.2470 [hep-ph] (cit. on pp. 27, 48).

[33] Frédéric A. Dreyer et al. "Recursive Soft Drop". In: *JHEP* 06 (2018), p. 093. DOI: 10.1007/JHEP06(2018)093. arXiv: 1804.03657 [hep-ph] (cit. on p. 27).

[34] Yacine Mehtar-Tani, Alba Soto-Ontoso, and Konrad Tywoniuk. "Dynamical grooming of QCD jets". In: *Phys. Rev. D* 101.3 (2020), p. 034004. DOI: 10.1103/PhysRevD.101.034004. arXiv: 1911.00375 [hep-ph] (cit. on p. 27).

[35] Raghunath Sahoo and Tapan Kumar Nayak. "Possible early universe signals in proton collisions at the Large Hadron Collider". In: *Curr. Sci.* 121 (2021), p. 1403. DOI: 10.18520/cs/v121/i11/1403-1408. arXiv: 2201.00202 [hep-ph] (cit. on p. 30).

[36] S. Durr et al. "Ab-Initio Determination of Light Hadron Masses". In: *Science* 322 (2008), pp. 1224–1227. DOI: 10.1126/science.1163233. arXiv: 0906.3599 [hep-lat] (cit. on p. 29).

[37] Y. Aoki et al. "The Order of the quantum chromodynamics transition predicted by the standard model of particle physics". In: *Nature* 443 (2006), pp. 675–678. DOI: 10.1038/nature05120. arXiv: hep-lat/0611014 (cit. on pp. 30, 32).

[38] R. Hagedorn. "Statistical thermodynamics of strong interactions at high-energies". In: *Nuovo Cim. Suppl.* 3 (1965), pp. 147–186 (cit. on p. 30).

[39] P. Kovtun, Dan T. Son, and Andrei O. Starinets. "Viscosity in strongly interacting quantum field theories from black hole physics". In: *Phys. Rev. Lett.* 94 (2005), p. 111601. DOI: 10.1103/PhysRevLett.94.111601. arXiv: hep-th/0405231 (cit. on p. 31).

[40] Elias R. Most et al. "Probing neutron-star matter in the lab: Similarities and differences between binary mergers and heavy-ion collisions". In: *Phys. Rev. D* 107.4 (2023), p. 043034. DOI: 10.1103/PhysRevD.107.043034. arXiv: 2201.13150 [nucl-th] (cit. on pp. 31, 32).

[41] G. Fiorella Burgio et al. "A Modern View of the Equation of State in Nuclear and Neutron Star Matter". In: *Symmetry* 13.3 (2021), p. 400. DOI: 10.3390/sym13030400 (cit. on p. 32).

[42] Bastian B. Brandt and Gergely Endrodi. "Reliability of Taylor expansions in QCD". In: *Phys. Rev. D* 99.1 (2019), p. 014518. DOI: 10.1103/PhysRevD.99.014518. arXiv: 1810.11045 [hep-lat] (cit. on p. 32).

[43] B. B. Brandt, G. Endrodi, and S. Schmalzbauer. "QCD phase diagram for nonzero isospin-asymmetry". In: *Phys. Rev. D* 97.5 (2018), p. 054514. DOI: 10.1103/PhysRevD.97.054514. arXiv: 1712.08190 [hep-lat] (cit. on p. 32).

[44] Debarati Chatterjee and Isaac Vidaña. "Do hyperons exist in the interior of neutron stars?" In: *Eur. Phys. J. A* 52.2 (2016), p. 29. DOI: 10.1140/epja/i2016-16029-x. arXiv: 1510.06306 [nucl-th] (cit. on p. 32).

[45] Ryan Abbott et al. "Lattice quantum chromodynamics at large isospin density". In: *Phys. Rev. D* 108.11 (2023), p. 114506. DOI: 10.1103/PhysRevD.108.114506. arXiv: 2307.15014 [hep-lat] (cit. on p. 32).

[46] Adam Bzdak et al. "Mapping the Phases of Quantum Chromodynamics with Beam Energy Scan". In: *Phys. Rept.* 853 (2020), pp. 1–87. DOI: 10.1016/j.physrep.2020.01.005. arXiv: 1906.00936 [nucl-th] (cit. on pp. 32, 33).

[47] Kari J. Eskola et al. "EPPS16: Nuclear parton distributions with LHC data". In: *Eur. Phys. J. C* 77.3 (2017), p. 163. DOI: 10.1140/epjc/s10052-017-4725-9. arXiv: 1612.05741 [hep-ph] (cit. on p. 33).

[48] C. Patrignani et al. "Review of Particle Physics". In: *Chin. Phys. C* 40.10 (2016), p. 100001. DOI: 10.1088/1674-1137/40/10/100001 (cit. on p. 33).

[49] B. Alver et al. "The PHOBOS Glauber Monte Carlo". In: (May 2008). arXiv: 0805.4411 [nucl-ex] (cit. on pp. 33, 34).

[50] Guillaume Falmagne. "The B_c^+ meson in heavy-ion collisions with the CMS detector". PhD thesis. Ecole Polytechnique, 2021 (cit. on p. 35).

[51] Aram Hayrapetyan et al. "Extracting the speed of sound in the strongly interacting matter created in ultrarelativistic lead-lead collisions at the LHC". In: (Jan. 2024). arXiv: 2401.06896 [nucl-ex] (cit. on pp. 35, 36).

[52] *lpthe.jussieu.fr*. https://www.lpthe.jussieu.fr/~soyez/data/talks/2020-11-SnowmassHI.pdf. [Accessed 18-06-2024] (cit. on p. 37).

[53] Jorge Casalderrey-Solana et al. "A hybrid strong/weak coupling approach to jet quenching". In: *JHEP* 10 (2014). [Erratum: JHEP 09 (2015) 175], p. 019. DOI: 10.1007/JHEP09(2015)175. arXiv: 1405.3864 [hep-ph] (cit. on pp. 38, 41, 180, 183).

[54] P. Caucal et al. "Vacuum-like jet fragmentation in a dense QCD medium". In: *Phys. Rev. Lett.* 120 (2018), p. 232001. DOI: 10.1103/PhysRevLett.120.232001. arXiv: 1801.09703 [hep-ph] (cit. on pp. 38, 45).

[55] Miklos Gyulassy and Michael Plumer. "Jet quenching in dense matter". In: *Phys. Lett. B* 243 (1990), p. 432. DOI: 10.1016/0370-2693(90)91409-5 (cit. on p. 38).

[56] R. Baier et al. "Radiative energy loss of high-energy quarks and gluons in a finite volume quark-gluon plasma". In: *Nucl. Phys. B* 483 (1997), p. 291. DOI: 10.1016/S0550-3213(96)00553-6. arXiv: hep-ph/9607355 (cit. on p. 38).

[57] R. Baier et al. "Radiative energy loss and p_T broadening of high-energy partons in nuclei". In: *Nucl. Phys. B* 484 (1997), p. 265. DOI: 10.1016/S0550-3213(96)00581-0. arXiv: hep-ph/9608322 (cit. on p. 38).

[58] Urs Achim Wiedemann. "Gluon radiation off hard quarks in a nuclear environment: Opacity expansion". In: *Nucl. Phys. B* 588 (2000), pp. 303–344. DOI: 10.1016/S0550-3213(00)00457-0. arXiv: hep-ph/0005129 (cit. on p. 38).

[59] M. Gyulassy, P. Levai, and I. Vitev. "Reaction operator approach to nonAbelian energy loss". In: *Nucl. Phys. B* 594 (2001), pp. 371–419. DOI: 10.1016/S0550-3213(00)00652-0. arXiv: nucl-th/0006010 (cit. on p. 38).

[60] B. G. Zakharov. "Radiative energy loss of high-energy quarks in finite size nuclear matter and quark - gluon plasma". In: *JETP Lett.* 65 (1997), pp. 615–620. DOI: 10.1134/1.567389. arXiv: hep-ph/9704255 (cit. on p. 38).

[61] L. D. Landau and I. Pomeranchuk. "Limits of applicability of the theory of bremsstrahlung electrons and pair production at high-energies". In: *Dokl. Akad. Nauk Ser. Fiz.* 92 (1953), pp. 535–536 (cit. on p. 39).

[62] Yacine Mehtar-Tani and Konrad Tywoniuk. "Improved opacity expansion for medium-induced parton splitting". In: *JHEP* 06 (2020), p. 187. DOI: 10.1007/JHEP06(2020)187. arXiv: 1910.02032 [hep-ph] (cit. on p. 39).

[63] Xabier Feal and Ricardo Vazquez. "Intensity of gluon bremsstrahlung in a finite plasma". In: *Phys. Rev. D* 98.7 (2018), p. 074029. DOI: 10.1103/PhysRevD.98.074029. arXiv: 1811.01591 [hep-ph] (cit. on p. 39).

[64] Carlota Andres, Liliana Apolinário, and Fabio Dominguez. "Medium-induced gluon radiation with full resummation of multiple scatterings for realistic parton-medium interactions". In: *JHEP* 07 (2020), p. 114. DOI: 10.1007/JHEP07(2020)114. arXiv: 2002.01517 [hep-ph] (cit. on p. 39).

[65] Jorge Casalderrey-Solana et al. "New picture of jet quenching dictated by color coherence". In: *Phys. Lett. B* 725 (2013), p. 357. DOI: 10.1016/j.physletb.2013.07.046. arXiv: 1210.7765 [hep-ph] (cit. on pp. 39, 40).

[66] J. Casalderrey-Solana and E. Iancu. "Interference effects in medium-induced gluon radiation". In: *JHEP* 08 (2011), p. 015. DOI: 10.1007/JHEP08(2011)015. arXiv: 1105.1760 [hep-ph] (cit. on p. 40).

[67] Yacine Mehtar-Tani, Carlos A. Salgado, and Konrad Tywoniuk. "Anti-angular ordering of gluon radiation in QCD media". In: *Phys. Rev. Lett.* 106 (2011), p. 122002. DOI: 10.1103/PhysRevLett.106.122002. arXiv: 1009.2965 [hep-ph] (cit. on p. 40).

[68] Daniel Pablos and Sergio Sanjurjo. "Color Coherence Effects in Dipole-Quark Scattering in the Soft Limit". In: (June 2024). arXiv: 2406.08550 [hep-ph] (cit. on p. 40).

[69] Francesco D'Eramo, Krishna Rajagopal, and Yi Yin. "Molière scattering in quark-gluon plasma: finding point-like scatterers in a liquid". In: *JHEP* 01 (2019), p. 172. DOI: 10.1007/JHEP01(2019)172. arXiv: 1808.03250 [hep-ph] (cit. on pp. 40, 41, 183).

[70] V. N. Gribov and L. N. Lipatov. "Deep inelastic e p scattering in perturbation theory". In: *Sov. J. Nucl. Phys.* 15 (1972), pp. 438–450 (cit. on p. 41).

[71] Guido Altarelli and G. Parisi. "Asymptotic Freedom in Parton Language". In: *Nucl. Phys. B* 126 (1977), pp. 298–318. DOI: 10.1016/0550-3213(77)90384-4 (cit. on p. 41).

[72] Yuri L. Dokshitzer. "Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics." In: *Sov. Phys. JETP* 46 (1977), pp. 641–653 (cit. on p. 41).

[73] Francesco D'Eramo et al. "Momentum broadening in weakly coupled quark-gluon plasma (with a view to finding the quasiparticles within liquid quark-gluon plasma)". In: *JHEP* 05 (2013), p. 031. DOI: 10.1007/JHEP05(2013)031. arXiv: 1211.1922 [hep-ph] (cit. on pp. 41, 183).

[74] Jorge Casalderrey-Solana et al. "Jet wake from linearized hydrodynamics". In: *JHEP* 05 (2021), p. 230. DOI: 10.1007/JHEP05(2021)230. arXiv: 2010.01140 [hep-ph] (cit. on pp. 41, 183).

[75] Albert M Sirunyan et al. "Jet Shapes of Isolated Photon-Tagged Jets in Pb-Pb and pp Collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV". In: *Phys. Rev. Lett.* 122 (2019), p. 152001. DOI: 10.1103/PhysRevLett.122.152001. arXiv: 1809.08602 [hep-ex] (cit. on pp. 42, 43, 50, 91, 92).

[76] Vardan Khachatryan et al. "Measurement of inclusive jet cross sections in pp and PbPb collisions at $\sqrt{s_{NN}} = 2.76$ TeV". In: *Phys. Rev. C* 96.1 (2017), p. 015202. DOI: 10.1103/PhysRevC.96.015202. arXiv: 1609.05383 [nucl-ex] (cit. on pp. 42, 43).

[77] Georges Aad et al. "Measurements of the suppression and correlations of dijets in Pb+Pb collisions at sNN=5.02 TeV". In: *Phys. Rev. C* 107.5 (2023). [Erratum: Phys.Rev.C 109, 029901 (2024)], p. 054908. DOI: 10.1103/PhysRevC.107.054908. arXiv: 2205.00682 [nucl-ex] (cit. on pp. 43, 44).

[78] Morad Aaboud et al. "Measurement of jet fragmentation in PbPb and pp collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV with the ATLAS detector". In: *Phys. Rev. C* 98 (2018), p. 024908. DOI: 10.1103/PhysRevC.98.024908. arXiv: 1805.05424 [nucl-ex] (cit. on pp. 46, 47).

[79] Shreyasi Acharya et al. "Medium modification of the shape of small-radius jets in central PbPb collisions at $\sqrt{s_{\mathrm{NN}}} = 2.76$ TeV". In: *JHEP* 10 (2018), p. 139. DOI:

10.1007/JHEP10(2018)139. arXiv: 1807.06854 [nucl-ex] (cit. on pp. 46, 47, 119).

[80] Albert M Sirunyan et al. "Measurement of the Splitting Function in pp and Pb-Pb Collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV". In: *Phys. Rev. Lett.* 120 (2018), p. 142302. DOI: 10.1103/PhysRevLett.120.142302. arXiv: 1708.09429 [nucl-ex] (cit. on pp. 47, 48).

[81] Shreyasi Acharya et al. "Measurement of the groomed jet radius and momentum splitting fraction in pp and PbPb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02\,\mathrm{TeV}$". In: *Phys. Rev. Lett.* 128 (2022), p. 102001. DOI: 10.1103/PhysRevLett.128.102001. arXiv: 2107.12984 [nucl-ex] (cit. on pp. 47, 48, 119, 187).

[82] Serguei Chatrchyan et al. "Observation and studies of jet quenching in PbPb collisions at nucleon-nucleon center-of-mass energy $\sqrt{s_{\mathrm{NN}}} = 2.76\,\mathrm{TeV}$". In: *Phys. Rev. C* 84 (2011), p. 024906. DOI: 10.1103/PhysRevC.84.024906. arXiv: 1102.1957 [nucl-ex] (cit. on pp. 47, 114, 119).

[83] Albert M Sirunyan et al. "First measurement of large area jet transverse momentum spectra in heavy ion collisions". In: *JHEP* 05 (2021), p. 284. DOI: 10.1007/JHEP05(2021)284. arXiv: 2102.13080 [hep-ex] (cit. on p. 47).

[84] Albert M Sirunyan et al. "Study of jet quenching with isolated-photon+jet correlations in PbPb and pp collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV". In: *Phys. Lett. B* 785 (2018), p. 14. DOI: 10.1016/j.physletb.2018.07.061. arXiv: 1711.09738 [nucl-ex] (cit. on pp. 47, 50, 51, 135).

[85] Jaroslav Adam et al. "Measurement of jet quenching with semi-inclusive hadron-jet distributions in central PbPb collisions at $\sqrt{s_{\mathrm{NN}}} = 2.76\,\mathrm{TeV}$". In: *JHEP* 09 (2015), p. 170. DOI: 10.1007/JHEP09(2015)170. arXiv: 1506.03984 [nucl-ex] (cit. on p. 47).

[86] Shreyasi Acharya et al. "Measurements of inclusive jet spectra in pp and central PbPb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02\,\mathrm{TeV}$". In: *Phys. Rev. C* 101 (2020), p. 034911. DOI: 10.1103/PhysRevC.101.034911. arXiv: 1909.09718 [nucl-ex] (cit. on p. 47).

[87] Shreyasi Acharya et al. "Exploration of jet substructure using iterative declustering in pp and PbPb collisions at LHC energies". In: *Phys. Lett. B* 802 (2020), p. 135227. DOI: 10.1016/j.physletb.2020.135227. arXiv: 1905.02512 [nucl-ex] (cit. on p. 47).

[88] Morad Aaboud et al. "Measurement of photon-jet transverse momentum correlations in 5.02 TeV PbPb and pp collisions with ATLAS". In: *Phys. Lett. B* 789 (2019), p. 167. DOI: 10.1016/j.physletb.2018.12.023. arXiv: 1809.07280 [nucl-ex] (cit. on pp. 47, 50).

[89] Morad Aaboud et al. "Comparison of fragmentation functions for jets dominated by light quarks and gluons from pp and PbPb collisions in ATLAS". In: *Phys. Rev. Lett.* 123 (2019), p. 042001. DOI: 10.1103/PhysRevLett.123.042001. arXiv: 1902.10007 [nucl-ex] (cit. on p. 47).

[90] Jaroslav Adam et al. "Measurement of groomed jet substructure observables in pp collisions at $\sqrt{s} = 200\,\mathrm{GeV}$ with STAR". In: *Phys. Lett. B* 811 (2020), p. 135846. DOI: 10.1016/j.physletb.2020.135846. arXiv: 2003.02114 [hep-ex] (cit. on pp. 47, 48).

[91] M. S. Abdallah et al. "Differential measurements of jet substructure and partonic energy loss in AuAu collisions at $\sqrt{s_{\mathrm{NN}}} = 200\,\mathrm{GeV}$". In: *Phys. Rev. C* 105 (2022), p. 044906. DOI: 10.1103/PhysRevC.105.044906. arXiv: 2109.09793 [nucl-ex] (cit. on pp. 47, 48).

[92] Liliana Apolinário, Yen-Jie Lee, and Michael Winn. "Heavy quarks and jets as probes of the QGP". In: *Prog. Part. Nucl. Phys.* 127 (2022), p. 103990. DOI: 10.1016/j.ppnp.2022.103990. arXiv: 2203.16352 [hep-ph] (cit. on p. 47).

[93] Leticia Cunqueiro and Anne M. Sickles. "Studying the QGP with jets at the LHC and RHIC". In: *Prog. Part. Nucl. Phys.* 124 (2022), p. 103940. DOI: 10.1016/j.ppnp.2022.103940. arXiv: 2110.14490 [nucl-ex] (cit. on p. 47).

[94] Andrew J. Larkoski et al. "Soft drop". In: *JHEP* 05 (2014), p. 146. DOI: 10.1007/JHEP05(2014)146. arXiv: 1402.2657 [hep-ph] (cit. on p. 48).

[95] Aram Hayrapetyan et al. "Girth and groomed radius of jets recoiling against isolated photons in lead-lead and proton-proton collisions at $\sqrt{s_{\mathrm{NN}}}$ = 5.02 TeV". In: *accepted by Phys. Lett. B* (May 2024). arXiv: 2405.02737 [nucl-ex] (cit. on pp. 49, 53, 98, 191).

[96] G. Aad et al. "Measurement of substructure-dependent jet suppression in PbPb collisions at 5.02 TeV with the ATLAS detector". In: *Phys. Rev. C* 107 (2023), p. 054909. DOI: 10.1103/PhysRevC.107.054909. arXiv: 2211.11470 [nucl-ex] (cit. on pp. 48, 52, 119, 187).

[97] Jasmine Brewer, José Guilherme Milhano, and Jesse Thaler. "Sorting out quenched jets". In: *Phys. Rev. Lett.* 122 (2019), p. 222301. DOI: 10.1103/PhysRevLett.122.222301. arXiv: 1812.05111 [hep-ph] (cit. on p. 49).

[98] Yi-Lun Du, Daniel Pablos, and Konrad Tywoniuk. "Deep learning jet modifications in heavy ion collisions". In: *JHEP* 21 (2020), p. 206. DOI: 10.1007/JHEP03(2021)206. arXiv: 2012.07797 [hep-ph] (cit. on p. 49).

[99] Jasmine Brewer, Quinn Brodsky, and Krishna Rajagopal. "Disentangling jet modification in jet simulations and in Z +jet data". In: *JHEP* 02 (2022), p. 175. DOI: 10.1007/JHEP02(2022)175. arXiv: 2110.13159 [hep-ph] (cit. on p. 49).

[100] Albert M Sirunyan et al. "The production of isolated photons in PbPb and pp collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02\,\mathrm{TeV}$". In: *JHEP* 07 (2020), p. 116. DOI: 10.1007/JHEP07(2020)116. arXiv: 2003.12797 [hep-ex] (cit. on pp. 50, 91, 95, 98, 99).

[101] Georges Aad et al. "Centrality, rapidity and transverse momentum dependence of isolated prompt photon production in lead-lead collisions at $\sqrt{s_{\mathrm{NN}}} = 2.76\,\mathrm{TeV}$ measured with the ATLAS detector". In: *Phys. Rev. C* 93 (2016), p. 034914. DOI: 10.1103/PhysRevC.93.034914. arXiv: 1506.08552 [hep-ex] (cit. on pp. 50, 91).

[102] Jaroslav Adam et al. "Direct photon production in PbPb collisions at $\sqrt{s_{\mathrm{NN}}} = 2.76\,\mathrm{TeV}$". In: *Phys. Lett. B* 754 (2016), p. 235. DOI: 10.1016/j.physletb.2016.01.020. arXiv: 1509.07324 [nucl-ex] (cit. on pp. 50, 91).

[103] L. Adamczyk et al. "Direct virtual photon production in AuAu collisions at $\sqrt{s_{\mathrm{NN}}} = 200\,\mathrm{GeV}$". In: *Phys. Lett. B* 770 (2017), p. 451. DOI: 10.1016/j.physletb.2017.04.050. arXiv: 1607.01447 [nucl-ex] (cit. on pp. 50, 91).

[104] S. S. Adler et al. "Centrality dependence of direct photon production in $\sqrt{s_{\mathrm{NN}}} = 200\,\mathrm{GeV}$ AuAu collisions". In: *Phys. Rev. Lett.* 94 (2005), p. 232301. DOI: 10.1103/PhysRevLett.94.232301. arXiv: nucl-ex/0503003 (cit. on pp. 50, 91).

[105] Serguei Chatrchyan et al. "Measurement of the differential cross section for isolated prompt photon production in pp collisions at 7 TeV". In: *Phys. Rev. D* 84 (2011), p. 052011. DOI: 10.1103/PhysRevD.84.052011. arXiv: 1108.2044 [hep-ex] (cit. on p. 50).

[106] Thomas Becher, Samuel Favrod, and Xiaofeng Xu. "QCD anatomy of photon isolation". In: *JHEP* 01 (2023), p. 005. DOI: 10.1007/JHEP01(2023)005. arXiv: 2208.01554 [hep-ph] (cit. on p. 50).

[107] Albert M Sirunyan et al. "Observation of medium-induced modifications of jet fragmentation in PbPb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02\,\mathrm{TeV}$ using isolated photon-tagged jets". In: *Phys. Rev. Lett.* 121 (2018), p. 242301. DOI: 10.1103/PhysRevLett.121.242301. arXiv: 1801.04895 [hep-ex] (cit. on p. 51).

[108] Georges Aad et al. "Comparison of inclusive and photon-tagged jet suppression in 5.02 TeV Pb+Pb collisions with ATLAS". In: *Phys. Lett. B* 846 (2023), p. 138154. DOI: 10.1016/j.physletb.2023.138154. arXiv: 2303.10090 [nucl-ex] (cit. on p. 52).

[109] "Comparison of inclusive and photon-tagged jet suppression in 5.02 TeV Pb+Pb collisions with ATLAS". In: (2022) (cit. on p. 52).

[110] T. Fazzini et al. "Electron Decay of the Pion". In: *Phys. Rev. Lett.* 1 (1958), pp. 247–249. DOI: 10.1103/PhysRevLett.1.247 (cit. on p. 55).

[111] "LHC Design Report Vol.1: The LHC Main Ring". In: (June 2004). Ed. by Oliver S. Bruning et al. DOI: 10.5170/CERN-2004-003-V-1 (cit. on p. 55).

[112] Esma Mobs. "The CERN accelerator complex - August 2018. Complexe des accélérateurs du CERN - Août 2018". In: (2018). General Photo. URL: https://cds.cern.ch/record/2636343 (cit. on p. 56).

[113] K. Aamodt et al. "The ALICE experiment at the CERN LHC". In: *JINST* 3 (2008), S08002. DOI: 10.1088/1748-0221/3/08/S08002 (cit. on p. 57).

[114] G. Aad et al. "The ATLAS Experiment at the CERN Large Hadron Collider". In: *JINST* 3 (2008), S08003. DOI: 10.1088/1748-0221/3/08/S08003 (cit. on p. 57).

[115] S. Chatrchyan et al. "The CMS experiment at the CERN LHC". In: *JINST* 3 (2008), S08004. DOI: 10.1088/1748-0221/3/08/S08004 (cit. on pp. 58, 63).

[116] A. Augusto Alves Jr. et al. "The LHCb Detector at the LHC". In: *JINST* 3 (2008), S08005. DOI: 10.1088/1748-0221/3/08/S08005 (cit. on p. 58).

[117] O. Adriani et al. "The LHCf detector at the CERN Large Hadron Collider". In: *JINST* 3 (2008), S08006. DOI: 10.1088/1748-0221/3/08/S08006 (cit. on p. 58).

[118] G. Anelli et al. "The TOTEM experiment at the CERN Large Hadron Collider". In: *JINST* 3 (2008), S08007. DOI: 10.1088/1748-0221/3/08/S08007 (cit. on p. 58).

[119] Malcolm Fairbairn and James L. Pinfold. "MoEDAL – a new light on the high-energy frontier". In: *Contemp. Phys.* 58.1 (2017), pp. 1–24. DOI: 10.1080/00107514.2016.1222649 (cit. on p. 58).

[120] Albert M Sirunyan et al. "Performance of the CMS Level-1 trigger in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$". In: *JINST* 15 (2020), P10017. DOI: 10.1088/1748-0221/15/10/P10017. arXiv: 2006.10165 [hep-ex] (cit. on pp. 59, 68, 82).

[121] Vardan Khachatryan et al. "The CMS trigger system". In: *JINST* 12 (2017), P01020. DOI: 10.1088/1748-0221/12/01/P01020. arXiv: 1609.02366 [physics.ins-det] (cit. on pp. 59, 82, 83).

[122] Albert M Sirunyan et al. "Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC". In: *JINST* 16 (2021), P05014. DOI: 10.1088/1748-0221/16/05/P05014. arXiv: 2012.06888 [hep-ex] (cit. on pp. 59, 75, 77, 78, 80, 82, 85, 86, 97, 117).

[123] A. M. Sirunyan et al. "Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$". In: *JINST* 13 (2018), P06015. DOI: 10.1088/1748-0221/13/06/P06015. arXiv: 1804.04528 [physics.ins-det] (cit. on p. 59).

[124] Serguei Chatrchyan et al. "Description and performance of track and primary-vertex reconstruction with the CMS tracker". In: *JINST* 9 (2014), P10009. DOI: 10.1088/1748-0221/9/10/P10009. arXiv: 1405.6569 [physics.ins-det] (cit. on pp. 59, 80).

[125] Albert M Sirunyan et al. "Precision luminosity measurement in proton-proton collisions at $\sqrt{s} = 13$ TeV in 2015 and 2016 at CMS". In: *Eur. Phys. J. C* 81 (2021), p. 800. DOI: 10.1140/epjc/s10052-021-09538-2. arXiv: 2104.01927 [hep-ex] (cit. on p. 59).

[126] *CMS luminosity measurement for the 2018 data-taking period at* $\sqrt{s} = 13$ *TeV*. CMS Physics Analysis Summary CMS-PAS-LUM-18-002. https://cds.cern.ch/record/2676164. 2019. URL: https://cds.cern.ch/record/2676164 (cit. on p. 59).

[127] CMS Collaboration. *Luminosity measurement in proton-proton collisions at 5.02 TeV in 2017 at CMS*. CMS Physics Analysis Summary CMS-PAS-LUM-19-001. http://cds.cern.ch/record/2765655. 2021. URL: http://cds.cern.ch/record/2765655 (cit. on p. 59).

[128] Tai Sakuma and Thomas McCauley. "Detector and Event Visualization with SketchUp at the CMS Experiment". In: *J. Phys. Conf. Ser.* 513 (2014). Ed. by D. L. Groep and D. Bonacorsi, p. 022032. DOI: 10.1088/1742-6596/513/2/022032. arXiv: 1311.4942 [physics.ins-det] (cit. on p. 59).

[129] Neutelings, Izaak. *Figures from TikZ.net*. (Retrieved May 28, 2024). 2021. URL: https://tikz.net/author/izaak/ (cit. on pp. 60, 61, 112).

[130] S Chatrchyan et al. "Precise Mapping of the Magnetic Field in the CMS Barrel Yoke using Cosmic Rays". In: *JINST* 5 (2010), T03021. DOI: 10.1088/1748-0221/5/03/T03021. arXiv: 0910.5530 [physics.ins-det] (cit. on p. 62).

[131] Dmitri Konstantinov. "Calibration and Performance of the CMS Electromagnetic Calorimeter in LHC Run 2". In: *PoS* ICHEP2020 (2021), p. 804. DOI: 10.22323/1.390.0804 (cit. on pp. 64, 65).

[132] "CMS Technical Design Report for the Phase 1 Upgrade of the Hadron Calorimeter". In: (Sept. 2012). Ed. by J. Mans et al. DOI: 10.2172/1151651 (cit. on p. 65).

[133] P. Paolucci et al. "CMS Resistive Plate Chamber overview, from the present system to the upgrade phase I". In: *PoS* RPC2012 (2012), p. 004. DOI: 10.1088/1748-0221/8/04/P04005. arXiv: 1209.1941 [physics.ins-det] (cit. on p. 66).

[134] S Chatrchyan et al. "Commissioning of the CMS High-Level Trigger with Cosmic Rays". In: *JINST* 5 (2010), T03005. DOI: 10.1088/1748-0221/5/03/T03005. arXiv: 0911.4889 [physics.ins-det] (cit. on p. 67).

[135] *Introduction to the CMS Trigger — cms-opendata-workshop.github.io*. https://cms-opendata-workshop.github.io/workshop2022-lesson-trigger/aio/index.html. [Accessed 01-07-2024] (cit. on p. 69).

[136] A. M. Sirunyan et al. "Particle-flow reconstruction and global event description with the CMS detector". In: *JINST* 12 (2017), P10003. DOI: 10.1088/1748-0221/12/10/P10003. arXiv: 1706.04965 [physics.ins-det] (cit. on pp. 72, 75, 77–79).

[137] Jonathan M. Butterworth, Guenther Dissertori, and Gavin P. Salam. "Hard Processes in Proton-Proton Collisions at the Large Hadron Collider". In: *Ann. Rev. Nucl. Part. Sci.* 62 (2012), pp. 387–405. DOI: 10.1146/annurev-nucl-102711-094913. arXiv: 1202.0583 [hep-ex] (cit. on p. 73).

[138] W. Adam et al. "Track reconstruction in the CMS tracker". In: (Dec. 2005) (cit. on p. 73).

[139] Emilio Meschi et al. "Electron Reconstruction in the CMS Electromagnetic Calorimeter". In: (June 2001) (cit. on pp. 73, 80, 81, 85).

[140] Wolfgang Adam et al. "Reconstruction of Electrons with the Gaussian-Sum Filter in the CMS Tracker at the LHC". In: (Jan. 2005) (cit. on p. 74).

[141] Vardan Khachatryan et al. "Performance of Photon Reconstruction and Identification with the CMS Detector in Proton-Proton Collisions at sqrt(s) = 8 TeV". In: *JINST* 10.08 (2015), P08010. DOI: 10.1088/1748-0221/10/08/P08010. arXiv: 1502.02702 [physics.ins-det] (cit. on pp. 75, 85, 107).

[142] Kaya Tatar, Ran Bi, Inna Kucher, Yeonju Go. "Photon reconstruction and performance in PbPb and pp collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02\,\mathrm{TeV}$ (Internal)". In: (2020). CMS-AN-18-314 v4. URL: https://cms.cern.ch/iCMS/user/noteinfo?cmsnoteid=CMS%20AN-2018/314 (cit. on pp. 85, 87, 88, 95, 98, 99).

[143] Helge Voss et al. "TMVA, the Toolkit for Multivariate Data Analysis with ROOT". In: *XIth International Workshop on Advanced Computing and Analysis Techniques in Physics Research (ACAT)*. [PoS(ACAT)040]. 2007, p. 40. DOI: 10.22323/1.050.0040. arXiv: physics/0703039 (cit. on pp. 87, 97, 117).

[144] David d'Enterria and Juan Rojo. "Quantitative constraints on the gluon distribution function in the proton from collider isolated-photon data". In: *Nucl. Phys. B* 860 (2012), p. 311. DOI: 10.1016/j.nuclphysb.2012.03.003. arXiv: 1202.1762 [hep-ph] (cit. on pp. 89–92).

[145] Shreyasi Acharya et al. "Measurement of the inclusive isolated photon production cross section in pp collisions at $\sqrt{s} = 7$ TeV". In: *Eur. Phys. J. C* 79.11 (2019), p. 896. DOI: 10.1140/epjc/s10052-019-7389-9. arXiv: 1906.01371 [nucl-ex] (cit. on p. 89).

[146] Shreyasi Acharya et al. "Measurement of the inclusive isolated-photon production cross section in pp collisions at $\sqrt{s}$ = 13 TeV". In: (July 2024). arXiv: 2407.01165 [hep-ex] (cit. on pp. 89, 90).

[147] Albert M Sirunyan et al. "Measurement of differential cross sections for inclusive isolated-photon and photon+jets production in proton-proton collisions at $\sqrt{s} = 13$ TeV". In: *Eur. Phys. J. C* 79 (2019), p. 20. DOI: 10.1140/epjc/s10052-018-6482-9. arXiv: 1807.00782 [hep-ex] (cit. on pp. 89, 90).

[148] Shreyasi Acharya et al. "Direct photon production at low transverse momentum in proton-proton collisions at $\sqrt{\mathbf{s}} = \mathbf{2.76}$ and 8 TeV". In: *Phys. Rev. C* 99.2 (2019), p. 024912. DOI: 10.1103/PhysRevC.99.024912. arXiv: 1803.09857 [nucl-ex] (cit. on pp. 89, 90).

[149] Georges Aad et al. "Measurement of the inclusive isolated-photon cross section in pp collisions at $\sqrt{s} = 13$ TeV using 36 fb^{-1} of ATLAS data". In: *JHEP* 10 (2019), p. 203. DOI: 10.1007/JHEP10(2019)203. arXiv: 1908.02746 [hep-ex] (cit. on p. 90).

[150] Shreyasi Acharya et al. "Neutral pion and η meson production at mid-rapidity in Pb-Pb collisions at $\sqrt{s_{NN}}$ = 2.76 TeV". In: *Phys. Rev. C* 98.4 (2018), p. 044901. DOI: 10.1103/PhysRevC.98.044901. arXiv: 1803.05490 [nucl-ex] (cit. on p. 92).

[151] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. "FastJet user manual". In: *Eur. Phys. J. C* 72 (2012), p. 1896. DOI: 10.1140/epjc/s10052-012-1896-2. arXiv: 1111.6097 [hep-ph] (cit. on p. 101).

[152] "Jet algorithms performance in 13 TeV data". In: (2017) (cit. on p. 102).

[153] Matthew Nguyen. "Jet Reconstruction with Particle Flow in Heavy-Ion Collisions with CMS". In: *J. Phys. G* 38 (2011). Ed. by Yves Schutz and Urs Achim

Wiedemann, p. 124151. DOI: 10.1088/0954-3899/38/12/124151. arXiv: 1107.0179 [nucl-ex] (cit. on p. 102).

[154] Betty Abelev et al. "Measurement of event background fluctuations for charged particle jet reconstruction in PbPb collisions at $\sqrt{s_{\mathrm{NN}}} = 2.76\,\mathrm{TeV}$". In: *JHEP* 03 (2012), p. 053. DOI: 10.1007/JHEP03(2012)053. arXiv: 1201.2423 [hep-ex] (cit. on p. 103).

[155] Doga Gulhan, Kurt Jung, Raghav K. E., Chris McGinn, Pawan K. N., Stephan O., and Marta V. "Jet Underlying-Event Subtraction and Reconstruction Performance at $\sqrt{s_{\mathrm{NN}}} = 5.02\,\mathrm{TeV}$ PbPb Collisions (Internal)". In: (). CMS-AN-16-069 v4. URL: https://cms.cern.ch/iCMS/user/noteinfo?cmsnoteid=CMS%20AN-2016/069 (cit. on pp. 103, 105).

[156] Olga E., Hallie C. T., Dhanush H., et al. "Jet properties in PbPb and pp collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02\,\mathrm{TeV}$ (Internal)". In: (). CMS-AN-16-298 v6. URL: https://cms.cern.ch/iCMS/user/noteinfo?cmsnoteid=CMS%20AN-2016/298 (cit. on pp. 103, 107).

[157] Olga Kodolova et al. "The performance of the jet identification and reconstruction in heavy ions collisions with CMS detector". In: *Eur. Phys. J. C* 50 (2007), p. 117. DOI: 10.1140/epjc/s10052-007-0223-9 (cit. on p. 103).

[158] Peter Berta et al. "Particle-level pileup subtraction for jets and jet shapes". In: *JHEP* 06 (2014), p. 092. DOI: 10.1007/JHEP06(2014)092. arXiv: 1403.3108 [hep-ex] (cit. on pp. 103, 104).

[159] Serguei Chatrchyan et al. "Determination of jet energy calibration and transverse momentum resolution in CMS". In: *JINST* 6 (2011), P11002. DOI: 10.1088/1748-0221/6/11/P11002. arXiv: 1107.4277 [physics.ins-det] (cit. on pp. 105, 107).

[160] Vardan Khachatryan et al. "Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV". In: *JINST* 12 (2017), P02014. DOI: 10.1088/1748-0221/12/02/P02014. arXiv: 1607.03663 [hep-ex] (cit. on pp. 105, 106, 108, 135).

[161] S. Agostinelli et al. "GEANT4—a simulation toolkit". In: *Nucl. Instrum. Meth. A* 506 (2003), p. 250. DOI: 10.1016/S0168-9002(03)01368-8 (cit. on pp. 106, 111).

[162] Serguei Chatrchyan et al. "Performance of CMS Muon Reconstruction in pp Collision Events at $\sqrt{s} = 7$ TeV". In: *JINST* 7 (2012), P10002. DOI: 10.1088/1748-0221/7/10/P10002. arXiv: 1206.4071 [physics.ins-det] (cit. on p. 107).

[163] Serguei Chatrchyan et al. "Energy Calibration and Resolution of the CMS Electromagnetic Calorimeter in pp Collisions at $\sqrt{s} = 7$ TeV". In: *JINST* 8 (2013), P09009. DOI: 10.1088/1748-0221/8/09/P09009. arXiv: 1306.2016 [hep-ex] (cit. on p. 107).

[164] Olga E., Hallie C. T., Dhanush H., et al. "Jet energy corrections for inclusive jets in PbPb and pp collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV PbPb Collisions (Internal)". In: (). CMS-AN-17-118 v2. URL: https://cms.cern.ch/iCMS/user/noteinfo?cmsnoteid=CMS%20AN-2017/118 (cit. on p. 107).

[165] Batoul Diab. "Fragmentation of jets containing a J/ψ meson in PbPb and pp collisions at 5 TeV with the CMS detector". PhD thesis. Saclay, 2020 (cit. on pp. 108, 113).

[166] Torbjörn Sjöstrand et al. "An Introduction to PYTHIA8.2". In: *Comput. Phys. Commun.* 191 (2015), p. 159. DOI: 10.1016/j.cpc.2015.01.024. arXiv: 1410.3012 [hep-ph] (cit. on p. 111).

[167] Torbjorn Sjöstrand. "The merging of jets". In: *Phys. Lett. B* 142 (1984), p. 420. DOI: 10.1016/0370-2693(84)91354-6 (cit. on p. 111).

[168] Albert M Sirunyan et al. "Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements". In: *Eur. Phys. J. C* 80 (2020), p. 4. DOI: 10.1140/epjc/s10052-019-7499-4. arXiv: 1903.12179 [hep-ex] (cit. on p. 111).

[169] Albert M Sirunyan et al. "Development and validation of HERWIG 7 tunes from CMS underlying-event measurements". In: *Eur. Phys. J. C* 81 (2021), p. 312. DOI: 10.1140/epjc/s10052-021-08949-5. arXiv: 2011.03422 [hep-ex] (cit. on pp. 111, 133).

[170] Stefan Gieseke, P. Stephens, and Bryan Webber. "New formalism for QCD parton showers". In: *JHEP* 12 (2003), p. 045. DOI: 10.1088/1126-6708/2003/12/045. arXiv: hep-ph/0310083 (cit. on p. 111).

[171] Richard D. Ball et al. "Parton distributions for the LHC Run II". In: *JHEP* 04 (2015), p. 040. DOI: 10.1007/JHEP04(2015)040. arXiv: 1410.8849 [hep-ph] (cit. on pp. 111, 180).

[172] Mrinal Dasgupta, Lorenzo Magnea, and Gavin P. Salam. "Nonperturbative QCD effects in jets at hadron colliders". In: *JHEP* 02 (2008), p. 055. DOI: 10.1088/1126-6708/2008/02/055. arXiv: 0712.3014 [hep-ph] (cit. on pp. 116, 177, 179).

[173] James Mulligan and Mateusz Ploskon. "Identifying groomed jet splittings in heavy ion collisions". In: *Phys. Rev. C* 102 (2020), p. 044913. DOI: 10.1103/PhysRevC.102.044913. arXiv: 2006.01812 [hep-ph] (cit. on p. 119).

[174] W. T. Giele, E. W. Nigel Glover, and David A. Kosower. "Jet investigations using the radial moment". In: *Phys. Rev. D* 57 (1998), p. 1878. DOI: 10.1103/PhysRevD.57.1878. arXiv: hep-ph/9706210 (cit. on p. 119).

[175] Raghav Kunnawalkam Elayavalli and Korinna Christine Zapp. "Medium response in JEWEL and its impact on jet shape observables in heavy ion collisions". In: *JHEP* 07 (2017), p. 141. DOI: 10.1007/JHEP07(2017)141. arXiv: 1707.01539 [hep-ph] (cit. on p. 119).

[176] Ren-Zhuo Wan et al. "Jet shape modification at LHC energies by JEWEL". In: *Chin. Phys. C* 43 (2019), p. 054110. DOI: 10.1088/1674-1137/43/5/054110. arXiv: 1812.10062 [hep-ph] (cit. on p. 119).

[177] G. D'Agostini. "A multidimensional unfolding method based on Bayes' theorem". In: *Nucl. Instrum. Meth. A* 362 (1995), p. 487. DOI: 10.1016/0168-9002(95)00274-X (cit. on p. 128).

[178] Tim Adye. "Unfolding algorithms and tests using RooUnfold". In: *PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding*. 2011, p. 313. DOI: 10.5170/CERN-2011-006.313. arXiv: 1105.1160 [physics.data-an] (cit. on p. 128).

[179] Johannes Bellm et al. "HERWIG7.0/HERWIG++3.0 release note". In: *Eur. Phys. J. C* 76 (2016), p. 196. DOI: 10.1140/epjc/s10052-016-4018-8. arXiv: 1512.01178 [hep-ph] (cit. on pp. 133, 179).

[180] Johannes Bellm et al. "HERWIG7.2 release note". In: *Eur. Phys. J. C* 80 (2020), p. 452. DOI: 10.1140/epjc/s10052-020-8011-x. arXiv: 1912.06509 [hep-ph] (cit. on pp. 133, 179).

[181] Aram Hayrapetyan et al. "Measurement of energy correlators inside jets and determination of the strong coupling $\alpha_S(m_Z)$". In: (Feb. 2024). arXiv: 2402.13864 [hep-ex] (cit. on p. 154).

[182] Aram Hayrapetyan et al. "Measurement of multidifferential cross sections for dijet production in proton-proton collisions at $\sqrt{s}$ = 13 TeV". In: (Dec. 2023). arXiv: 2312.16669 [hep-ex] (cit. on p. 165).

[183] Walter T. Giele, David A. Kosower, and Peter Z. Skands. "A simple shower and matching algorithm". In: *Phys. Rev. D* 78 (2008), p. 014026. DOI: 10.1103/PhysRevD.78.014026. arXiv: 0707.3652 [hep-ph] (cit. on p. 179).

[184] Stefan Höche and Stefan Prestel. "The midpoint between dipole and parton showers". In: *Eur. Phys. J. C* 75 (2015), p. 461. DOI: 10.1140/epjc/s10052-015-3684-2. arXiv: 1506.05057 [hep-ph] (cit. on p. 179).

[185] A. Gehrmann-De Ridder, T. Gehrmann, and E. W. Nigel Glover. "Antenna subtraction at NNLO". In: *JHEP* 09 (2005), p. 056. DOI: 10.1088/1126-6708/2005/09/056. arXiv: hep-ph/0505111 (cit. on p. 179).

[186] Helen Brooks, Christian T. Preuss, and Peter Skands. "Sector showers for hadron collisions". In: *JHEP* 07 (2020), p. 032. DOI: 10.1007/JHEP07(2020)032. arXiv: 2003.00702 [hep-ph] (cit. on p. 179).

[187] S. Catani and M. H. Seymour. "A general algorithm for calculating jet cross sections in NLO QCD". In: *Nucl. Phys. B* 485 (1997). [Erratum: doi:10.1016/S0550-3213(98)81022-5], p. 291. DOI: 10.1016/S0550-3213(96)00589-5. arXiv: hep-ph/9605323 (cit. on p. 179).

[188] Peter Skands, Stefano Carrazza, and Juan Rojo. "Tuning PYTHIA8.1: the Monash 2013 tune". In: *Eur. Phys. J. C* 74 (2014), p. 3024. DOI: 10.1140/epjc/s10052-014-3024-y. arXiv: 1404.5630 [hep-ph] (cit. on p. 180).

[189] Zachary Hulcher, Daniel Pablos, and Krishna Rajagopal. "Resolution effects in the hybrid strong/weak coupling model". In: *JHEP* 03 (2018), p. 010. DOI: 10.1007/JHEP03(2018)010. arXiv: 1707.05245 [hep-ph] (cit. on p. 183).

ECOLE DOCTORALE

Titre : Mesures de la sous-structure des jets marqués par des photons dans les collisions pp et PbPb avec le détecteur CMS

Mots clés : CERN,CMS,sous-structure des jets,LHC, bosons électrofaibles, plasma quarks gluons

Résumé : Cette thèse présente les mesures de la circonférence g et du groomed des jets (R_{g}) reculant contre des photons isolés dans des collisions plomb-plomb (PbPb) et proton-proton (pp) au LHC à 5.02 TeV à l'aide des données collectées par l'expérience CMS. Ces mesures permettent de mieux comprendre les propriétés du plasma de quarks et de gluons (QGP), un milieu chaud et dense créé dans les collisions PbPb. Les modifications des jets dans le milieu QGP sont étudiées par rapport à une référence fournie par les événements de collisions pp. Les observables g et R_{g} fournissent une mesure quantitative de l'étroitesse ou de la largeur de la structure du jet. Les mesures de ces observables dans les événements de jet inclusif ont montré une sous-structure de jet plus étroite dans les collisions d'ions lourds par rapport aux collisions pp pour les jets reconstruits avec le même moment transverse de jet $p_{\mathrm{T}}^{\mathrm{jet}}$ dans les mesures précédentes. La mesure de ces observables dans les événements de jets marqués par des photons fournit une interprétation complémentaire de la perte d'énergie des jets dans le milieu QGP en considérant le moment transverse du photon p_{T}^{γ} comme une approximation du moment transverse du parton qui initie la gerbe de jets.
Les événements doivent comporter un photon avec un moment transverse $p_{\mathrm{T}}^{\gamma} > 100\,\mathrm{GeV}$ et au moins un jet dos à dos en azimut par rapport au photon et avec un moment transverse $p_{\mathrm{T}}^{\mathrm{jet}}$ tel que $p_{\mathrm{T}}^{\mathrm{jet}}/p_{\mathrm{T}}^{\gamma} > 0.4$. Les distributions R_{g} et g mesurées sont dépliées au niveau des particules, ce qui facilite la comparaison entre les résultats PbPb et pp et avec les prédictions théoriques. On constate que les jets avec $p_{\mathrm{T}}^{\mathrm{jet}}/p_{\mathrm{T}}^{\gamma} > 0.8$, c'est-à-dire ceux qui équilibrent étroitement le photon p_{T}^{γ}, sont plus étroits dans les collisions PbPb que dans les collisions pp. L'assouplissement de la sélection pour inclure les jets avec $p_{\mathrm{T}}^{\mathrm{jet}}/p_{\mathrm{T}}^{\gamma} > 0.4$ réduit le rétrécissement de la structure angulaire des jets dans PbPb par rapport à la référence pp. Ceci montre que les effets de biais de sélection associés à la perte d'énergie des jets jouent un rôle important dans l'interprétation des mesures de la sous-structure des jets.

Title : Photon tagged jet substructure measurements in pp and PbPb collisions with the CMS detector

Keywords : CERN, CMS, jet substructrure, LHC, electroweak bosons, quark-gluon plasma

Abstract : Measurements of the girth g and groomed radius of jets (R_{g}) recoiling against isolated photons in lead-lead (PbPb) and proton-proton (pp) collisions at the LHC at 5.02 TeV using the data collected at the CMS experiment are presented in this thesis. These measurements provide insight into the properties of the Quark Gluon Plasma (QGP), a hot and dense medium created in PbPb collisions. Modifications of jets in the QGP medium are studied relative to a reference provided by pp collision events. The observables g and R_{g} provide a quantitative measure of how narrow or broad the structure of the jet is. Measurements of these in inclusive jet events have shown a narrower jet substructure in heavy ion collisions compared with pp collisions for jets reconstructed with the same jet transverse momentum $p_{\mathrm{T}}^{\mathrm{jet}}$ in previous measurements. A measurement of these observables in photon-tagged jet events provides a complementary interpretation about the jet energy loss in the QGP medium by considering the photon transverse momentum p_{T}^{γ} as a proxy for the p_{T} of the parton that initiates the jet shower.
Events are required to have a photon with transverse momentum $p_{\mathrm{T}}^{\gamma} > 100\,\mathrm{GeV}$ and at least one jet back-to-back in azimuth with respect to the photon and with transverse momentum $p_{\mathrm{T}}^{\mathrm{jet}}$ such that $p_{\mathrm{T}}^{\mathrm{jet}}/p_{\mathrm{T}}^{\gamma} > 0.4$. The measured R_{g} and g distributions are unfolded to the particle level, which facilitates the comparison between the PbPb and pp results and with theoretical predictions. It is found that jets with $p_{\mathrm{T}}^{\mathrm{jet}}/p_{\mathrm{T}}^{\gamma} > 0.8$, i.e., those that closely balance the photon p_{T}^{γ}, are narrower in PbPb than in pp collisions. Relaxing the selection to include jets with $p_{\mathrm{T}}^{\mathrm{jet}}/p_{\mathrm{T}}^{\gamma} > 0.4$ reduces the narrowing of the angular structure of jets in PbPb relative to the pp reference. This shows that selection bias effects associated with jet energy loss play an important role in the interpretation of jet substructure measurements.

Institut Polytechnique de Paris
91120 Palaiseau, France